

TREES IN RANDOM SPARSE GRAPHS WITH A GIVEN DEGREE SEQUENCE

BEHZAD MEHRDAD

ABSTRACT. Let \mathbb{G}^D be the set of graphs $G(V, E)$ with $|V| = n$, and the degree sequence equal to $D = (d_1, d_2, \dots, d_n)$. In addition, for $\frac{1}{2} < a < 1$, we define the set of graphs with an almost given degree sequence \bar{D} as follows,

$$\mathbb{G}_a^D := \cup \mathbb{G}^{\bar{D}},$$

where the union is over all degree sequences \bar{D} such that, for $1 \leq i \leq n$, we have $|d_i - \bar{d}_i| < d_i^a$.

Now, if we chose random graphs $\mathcal{G}_{\mathbf{g}}(D)$ and $\mathcal{G}_{\mathbf{a}}(D)$ uniformly out of the sets \mathbb{G}^D and \mathbb{G}_a^D , respectively, what do they look like? This has been studied when $\mathcal{G}_{\mathbf{g}}(D)$ is a dense graph, i.e. $|E| = \Theta(n^2)$, in the sense of graphons, or when $\mathcal{G}_{\mathbf{g}}(D)$ is very sparse, i.e. $d_n^2 = o(|E|)$. In the case of sparse graphs with an almost given degree sequence, we investigate this question, and give the finite tree subgraph structure of $\mathcal{G}_{\mathbf{a}}(D)$ under some mild conditions. For the random graph $\mathcal{G}_{\mathbf{g}}(D)$ with a given degree sequence, we re-derive the finite tree structure in dense and very sparse cases to give a continuous picture.

Moreover, for a pair of vectors $(D_1, D_2) \in \mathbb{Z}^{n_1} \times \mathbb{Z}^{n_2}$, we let $\mathcal{G}_{\mathbf{b}}(D_1, D_2)$ be the random bipartite graph that is chosen uniformly out of the set \mathbb{G}^{D_1, D_2} , where \mathbb{G}^{D_1, D_2} is the set of all bipartite graphs with the degree sequence (D_1, D_2) . We are able to show the result for $\mathcal{G}_{\mathbf{b}}(D_1, D_2)$ without any further conditions.

CONTENTS

1. Introduction	2
1.1. Graphs with a given degree sequence.	2
1.2. Dense graphs:	2
1.3. Very sparse graphs:	3
1.4. Bipartite graphs:	3
1.5. A little bit of motivation:	3
2. Main Results.	4
2.1. Maximum entropy and the independent ensemble.	4
2.2. Ordered trees and their B- function.	5
2.3. Graphs with an almost given degree sequence:	6
2.4. Graphs with a given degree sequence.	8
2.5. Dense graphs with a given degree sequence:	9
2.6. Very sparse graphs with a given degree sequence:	10
2.7. Bipartite graphs with an exact degree sequence:	11
2.8. Open questions	13
3. Proofs	13
3.1. Graph with an almost given degree sequence.	13
3.2. A lower bound	26

3.3. Graphs with a given degree sequence.	34
3.4. Dense graphs (Definition 2.21).	35
3.5. Very Sparse graphs (Definition 2.24).	36
3.6. Bipartite graphs (proof of Theorem 2.30).	40
4. Acknowledgements	41
Appendix A. A concentration inequality.	41
Appendix B. Regularity of r_i s.	43
References	47

1. INTRODUCTION

1.1. Graphs with a given degree sequence. Let $D = (d_1, d_2, \dots, d_n)$ be a finite sequence of positive integers, such that $1 \leq d_1 \leq \dots \leq d_n$, and $M := \sum_{i=1}^n d_i$ is even. In addition, let \mathbb{G}_n be the set of all simple graphs (undirected, with no loops or multiple edges) with n vertices. For $G \in \mathbb{G}_n$, $\mathcal{V}(G)$ will be the set of its vertices indexed by $\{v_1, v_2, \dots, v_n\}$ and $\mathcal{E}(G)$ is its set of edges $e := \langle v_i, v_j \rangle$. We say that G is a graph with the given degree sequence D , if the degree of a vertex v_i is d_i . It is evident that the total number of edges $|\mathcal{E}(G)| = \frac{1}{2}M$. We denote the set of all such simple graphs by $\mathbb{G}^D \subset \mathbb{G}_n$.

A random graph $\mathcal{G}_{\mathbf{g}}(D)$ with the given degree sequence D is the one that is uniformly chosen from \mathbb{G}^D . Now, what does the random graph $\mathcal{G}_{\mathbf{g}}(D)$ look like? Researchers have studied this problem extensively, along with other properties of graphs with a given degree sequence. Before we list a couple of them here, we state some notations.

In this paper and for two real functions $f(n)$ and $g(n)$, the notations $g = \Theta(f)$, $g = O(f)$ and $g = o(f)$, as n goes to infinity, mean that there exists a real number C such that $\limsup_{n \rightarrow \infty} \frac{|f|}{|g|} + \frac{|g|}{|f|} \leq C$, $\limsup_{n \rightarrow \infty} \frac{|f|}{|g|} \leq C$ and $\lim_{n \rightarrow \infty} \frac{|f|}{|g|} = 0$, respectively.

1.2. Dense graphs: Dense graphs are those graphs $G \in \mathbb{G}_n$ that $|\mathcal{E}(G)| = \Theta(n^2)$. In [1], Barvinok and Hartigan studied the structure of dense graphs with a given degree sequence. They showed the relation between the maximum entropy function and the number of such graphs. Under mild conditions (δ -tameness), they found the asymptotic behavior of the number of graphs with a given degree sequence D .

Although Barvinok and Hartigan provide an exact formula, it is difficult to touch it. There are also some other approaches to the counting problem which only work in certain regimes. In [19], McKay and Wormald considered the case of graphs with nearly constant degree d , $|d_i - d| = O(n^{1/2})$, using a multidimensional saddle-point method. The enumeration of graphs with a given degree sequence may lead to finding the probabilities of subgraphs of random graph $\mathcal{G}_{\mathbf{g}}(D)$. Greenhill and McKay studied that in [29] for various regimes. Also, see McKay [28] for a detailed survey of that subject.

Another approach toward graphs with a given degree sequence is through graph limits. Recently Lovász and Szegedy introduced, in [23], a notion of graph limits called graphons. This has been developed further by Borgs et al [8, 10, 9]. In regard to graphs with a given degree sequence, Chatterjee et al [6] showed that sequences of such graphs have graph limits, in the sense of *graphons*, if their degree sequences

converge to a degree function which satisfies the Erdős-Gallai condition for graph limits.

1.3. Very sparse graphs: Different regimes of very sparse graphs, $d_{max}^2 = d_n^2 = o(M)$, were studied a long time ago by McKay, [27] and [26]. The condition allows us to compute the number of graphs via inclusion-exclusion and switching method. McKay and Wormald came back to this problem in [30] with a less restrictive condition. Recently Gao et al [18] investigated the probability of subgraphs of a random graph with a given degree sequence in this regime. Look at [17] for more information about subgraphs of random graphs.

1.4. Bipartite graphs: A bipartite graph is a graph with two set of vertices, where there are no edges with both ends in the same set. The adjacency matrix for a simple bipartite graph with a given degree sequence is a matrix with 0-1 entries with given row and column sums. Barvinok, in [2, 1], studied these matrices in the dense case. Barvinok and Hartigan generalized this in [4, 5] for matrices with non-negative integer entries. Like the case of usual graphs, they showed the relation of the number of bipartite graphs to entropy function. Look at [3] for a survey on the subject.

Canfield et al [11] derived a practical formula for matrices with 0-1 entries and nearly constant row and column sums. Canfield and McKay, in [12], took a look at matrices with positive integer entries as well. Also, see [20] for similar results in sparse bipartite graphs.

1.5. A little bit of motivation: Following the work of Lovász and Szegedy, many tried to extend the notion of graphons to the sparse graphs. Look at [7] for a survey of attempts to define a notion of limit in the sparse case. Here, we take another look at the subgraph counting metric. Let us recall the homomorphism density from page 2 of Lovász and Szegedy [23], which is

$$t(F, G) := \frac{\text{hom}(F, G)}{|\mathcal{V}(G)|^{|\mathcal{V}(F)|}}.$$

In addition, graphs G_n are said to be Cauchy in subgraph-counting metric if the sequence of numbers $t(F, G_n)$ are Cauchy for every finite graph F . So, we use a uniform normalization, i.e. $|\mathcal{V}(G)|^{|\mathcal{V}(F)|}$, for all embeddings of F into G . However, we believe that the normalization should be local and depend on the embedding. We try to justify that throughout this paper.

Although we will not provide a metric for sparse graphs, we make a few observations in sparse random graphs with a given degree sequence. We adopt the method in [6] that compares the random graph $\mathcal{G}_{\mathbf{g}}(D)$ with a random graph $\tilde{\mathcal{G}}(D)$, with independent Bernoulli random edges. By extending that work to sparse graphs, we obtain the correct normalization for counting the subgraph F , where F is a tree. We leave the case the counting of subgraphs with loops open, since this problem has not been completely understood even in the case of random models with independent edges like Erdős-Rényi graph. For more discussion, look at [22], [21], and [14]).

In this paper, we first introduce a modified version of graphs with the given degree sequence D , which we call graphs with an almost given degree sequence D . Then, under some mild conditions, we find the distribution of finite trees in this model. Second, we go back to our original problem and deal with random graph

$\mathcal{G}_{\mathbf{g}}(D)$. In addition, we apply our method to dense, bipartite and very sparse, i.e. $d_n^2 = o(|\mathcal{E}(G)|)$, random graphs.

Although we need some mild conditions in most of our theorems, we show that our results holds in full generality for bipartite random graphs. So we believe that the same is true for general non-bipartite graphs with a given degree sequence D . In the end, to the best of our knowledge, the method developed here is new and works for a wider range of graphs, from very sparse to dense graphs.

We begin the next section with some notation that is needed for the rest of the paper.

2. MAIN RESULTS.

Let us start this section by stating a few notations. Throughout this section $c_k > 0$ plays the role of a general constant that only depends on k . Now we provide the definitions of our independent model and ordered subgraphs.

2.1. Maximum entropy and the independent ensemble. We let \mathbb{P}^D be the set of all positive x_{ij} satisfying

$$\sum_{j: j \neq i}^n x_{ij} = d_i.$$

Therefore, \mathbb{P}^D is a polytope in \mathbb{R}^N , where N is $\binom{n}{2}$. Now, define the entropy function as follows,

$$(2.1) \quad H_1(x) := \sum_{i < j} H(x_{ij}), \text{ where } H(x) = -x \ln(x) - (1-x) \ln(1-x),$$

for $x = (x_{ij}) \in \mathbb{R}^N$. We state a proposition that describes the necessary and sufficient condition for \mathbb{P}^D to have a non-empty interior.

Proposition 2.1. *The polytope \mathbb{P}^D has a non-empty interior if, and only if, the degree sequence satisfies the strict Erdős- Gallai conditions,*

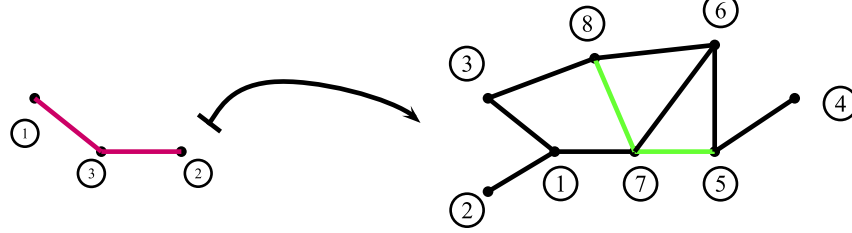
$$(2.2) \quad \sum_{i=n-k+1}^n d_i < k(k-1) + \sum_{i=1}^{n-k} \min\{k, d_i\}, \text{ for } 1 \leq k \leq n.$$

Remark 2.2. If we turn the strict inequalities in (2.2) into “less than or equal to”, then we obtain the well-known Erdős- Gallai criterion Erdős and Gallai [16]. (See Mahadev and Peled [24] for extensive discussions.)

Now, if \mathbb{P}^D has a non-empty interior, then the function $H_1(x)$ attains its maximum at a unique point, since it is strictly concave. Denote that maximum by $\tilde{\mathbf{p}} = (\tilde{p}_{ij}) \in \mathbb{R}^N$, and define $\tilde{\mathcal{G}}(D)$ as a random graph with independent Bernoulli random edges with parameters \tilde{p}_{ij} . From the definition of the $\tilde{\mathbf{p}}$ and for each $i \in [n] := \{1, \dots, n\}$, we see that the average degree of vertex i in $\tilde{\mathcal{G}}(D)$ is d_i . In other words,

$$(2.3) \quad \sum_{j: j \neq i}^n E \left[\mathbf{1}_{\langle i, j \rangle} \left(\tilde{\mathcal{G}}(D) \right) \right] = \sum_{j: j \neq i}^n \tilde{p}_{ij} = d_i,$$

FIGURE 2.1. An ordered tree



The left graph is a labeled tree $T \in \mathbb{T}^3$ with 2 edges. The green graph on the right is the image of T under a map s that takes 1, 3 and 2 to 8, 7 and 5 respectively.

Hence, the green graph is an ordered tree (s, T) that sits inside a graph with degree sequence $D = (3, 1, 2, 1, 3, 3, 4, 3)$. The corresponding B- function for (s, T) is $d_8^0 \cdot d_7^2 \cdot d_5^0 = 4^2$.

where the indicator function $\mathbf{1}_{\langle i, j \rangle}(\tilde{\mathcal{G}}(D))$ is 1 if $\langle i, j \rangle \in \tilde{\mathcal{G}}(D)$ and 0 otherwise. We also use “ \sim ” for parameters of $\tilde{\mathcal{G}}(D)$, wherever possible.

Definition 2.3. A vector of positive integers D is a *strict graphic sequence* if the vector D satisfies the strict Erdős- Gallai conditions (2.2).

Remark 2.4. Definition 2.3 means that \mathbb{P}^D has a non-empty interior, which in turn implies that the maximum entropy $\tilde{\mathbf{p}} = (\tilde{p}_{ij})$ exists.

2.2. Ordered trees and their B- function. Let \mathbb{T}^k be the set of all trees with a finite number k of vertices. The famous Cayley’s Theorem states that there are k^{k-2} of such trees, and for a proof of it, check [31]. For a tree $T \in \mathbb{T}^k$, we look at maps $s : \mathcal{V}(T) \rightarrow \mathcal{V}(G)$, which map the vertices of T into distinct vertices of \mathbb{G}_n . There are $n(n-1) \cdots (n-k+1)$ of them, and let us call the set of all such maps \mathbb{S}_n^k , i.e.

$$\mathbb{S}_n^k = \{s : [k] \rightarrow [n] \mid s \text{ is } 1-1\}.$$

In addition, an ordered tree is a pair $(s, T) \in \mathbb{S}_n^k \times \mathbb{T}^k$. For instance, if $k = 2$, $\mathbb{S}_n^2 \times \mathbb{T}^2$ is the set of directed edges on vertices of $[n] = \{1, \dots, n\}$. (We drop the n in the index of \mathbb{S}_n^k , whenever the dependency on n is understood.)

Definition 2.5. We let $(s, T) \in \mathbb{S}_n^k \times \mathbb{T}^k$ be an ordered tree. For each vertex $u \in \mathcal{V}(T)$, its degree in the tree T is denoted by b_u . The B - function is defined as

$$(2.4) \quad \psi(s, T, D) = \prod_{u \in \mathcal{V}(T)} d_{s(u)}^{b_u-1}.$$

In addition, we denote $\psi(s, T, D(G))$ by $\psi(s, T, G)$, where $D(G)$ is the degree sequence of G .

Remark 2.6. Let us consider a permutation π on numbers 1 through k . We observe that

$$\psi(s, T, D) = \psi(s \circ \pi^{-1}, \pi \circ T, D).$$

Hence, we get $k!$ distinct ordered trees with the same B -function.

2.3. Graphs with an almost given degree sequence: We recall from the beginning of this paper that $D = (d_1, \dots, d_n)$, and $d_1 \leq \dots \leq d_n$, and $M = \sum_{i=1}^n d_i$. We let S_k be the sum of biggest d_k elements of D , or $S_k := \sum d_i$, where the sum is over the set $\{n - d_k, \dots, n\}$. In addition, we define $\ell(D)$ as the maximum positive integer that $S_\ell \leq \frac{M}{2}$, i.e.

$$(2.5) \quad \ell(D) := \max \left\{ k \in [n] \mid S_k \leq \frac{M}{2} \right\}.$$

Assumption 2.7. Let us assume that, for some numbers $\epsilon > 0$ and $\nu > 0$,

- (1) the vector D satisfies the strict Erdős- Gallai conditions (2.2).
- (2) the number $M = \sum_{i=1}^n d_i$ is even and $n^{1+\epsilon} \leq M$.
- (3) and, for the function $\ell(D)$ as in (2.5),

$$(2.6) \quad \sqrt{\frac{n}{M}} (d_n - d_{\ell(D)} + 1) < n^{-\nu}.$$

Definition 2.8. In particular, we say that a vector D is a strict graphic sequence of type (ϵ, ν) , if it satisfies all of the above conditions, and is a strict graphic sequence of type ϵ , if D only satisfies the first two conditions.

Remark 2.9. Let us see a couple of examples to understand the term $d_n - d_{\ell(D)}$ and the above conditions. Suppose that $d_n^2 < \frac{M}{2}$ (particularly, this is particularly the case when $d_n^2 = o(M)$). Since d_n is the maximum element of D , $S_n = \sum_{i=1}^n d_{n+1-i}$ is less than d_n^2 , which means $\ell(D)$ is n . Hence, $d_n - d_{\ell(D)} = 0$.

In another example, let D be a sequence such that $2n^\alpha < d_i < n^\beta$, for $1 \leq i \leq n$, and $\beta < \frac{1+\alpha}{2}$. Again, S_n is less than d_n^2 , so $S_n < n^{2\beta} < n^{1+\alpha} < M/2$. In both examples, we get an upper bound of $\sqrt{\frac{n}{M}}$ for Eq. (2.6). In particular, our sequence is of type (ϵ, ν) for any ν smaller than $\epsilon/2$.

Let us pick a positive number a such that $\frac{1}{2} < a < 1$. Then, we define the set of graphs with an almost given degree sequence D as follows,

$$\mathbb{G}_a^D := \cup \mathbb{G}^{\bar{D}},$$

where the union is over all degree sequences \bar{D} such that, for $1 \leq i \leq n$, we have $|d_i - \bar{d}_i| < d_i^a$. Let a random graph with an *almost* given degree sequence D , $\mathcal{G}_a(a, D)$, be a random graph that is uniformly chosen from the set \mathbb{G}_a^D .

Definition 2.10. We define probabilities $\mathbf{p}_a(s, T)$ and $\tilde{\mathbf{p}}(s, T)$ as

$$E[\mathbf{1}_s(T, \mathcal{G}_a(a, D))] \text{ and } E[\mathbf{1}_s(T, \tilde{\mathcal{G}}(D))],$$

respectively, where

$$\mathbf{1}_s(T, G) = \prod_{\langle u_1, u_2 \rangle \in \mathcal{E}(T)} \mathbf{1}_{\langle s(u_1), s(u_2) \rangle \in \mathcal{E}(G)}(G),$$

$\mathcal{G}_a(a, D)$ is as above, and $\tilde{\mathcal{G}}(D)$ is defined in Section 2.1. We dropped the dependency of $\mathbf{p}_a(s, T)$ on D and a , as well as the dependency of $\tilde{\mathbf{p}}(s, T)$ on D and $\tilde{\mathbf{p}}$, the maximum entropy, for the simplicity of our notations.

Theorem 2.11. *Suppose that the vector D is a strict graphic sequence of type (ε, ν) as in Definition 2.3. We define,*

$$(2.7) \quad L_{\mathbf{a}}(a, k, D) := \frac{1}{M} \sum_{(s, T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} |\mathbf{p}_{\mathbf{a}}(s, T) - \tilde{\mathbf{p}}(s, T)|,$$

where \mathbb{T}^k , \mathbb{S}_n^k and $\psi(s, T, D)$ are defined in Section 2.2, and $\mathbf{p}_{\mathbf{a}}(s, T)$ and $\tilde{\mathbf{p}}(s, T)$ are the same as in Definition 2.10. If $\frac{1}{2} < a < \frac{\nu}{4} + \frac{1}{2}$, then

$$(2.8) \quad L_{\mathbf{a}}(a, k, D) \leq c_k \cdot n^{a - \frac{\nu+1}{2}} = o(1).$$

Remark 2.12. Recall from Remark 2.6 that $\psi(s, T, D)$ is invariant under the action of a permutation on the labels of an ordered tree (s, T) . In addition, the space $\mathbb{S}^k \times \mathbb{T}^k$ is a multi covering of

$$\mathcal{T}(k, n) := \cup_{T \in \mathbb{T}^k} \text{hom}(T, K_n)$$

with $k!$ layers, where K_n is the complete graph with n vertices. Hence, if we take the sum over $\mathcal{T}(k, n)$ instead of $\mathbb{S}_n^k \times \mathbb{T}^k$, we get $\frac{L_{\mathbf{a}}(a, k, D)}{k!}$, which is still $o(1)$. Although the set $\mathbb{S}_n^k \times \mathbb{T}^k$ requires us to over-count objects, it also brings us symmetry and that makes it easier to deal with.

Remark 2.13. For the examples in Remark 2.9, $n^{-\nu}$ is $\sqrt{\frac{n}{M}}$ and the bound in Eq. (2.8) becomes $O\left(\left(\frac{n}{M}\right)^{1/4} n^{a - \frac{1}{2}}\right)$. We believe that this is the correct bound on $L_{\mathbf{a}}(a, k, D)$, and also, we believe the last condition in Definition 2.3 is not necessary. We come back to this matter later in the proof section (Conjecture 3.13 and Remark 3.14).

Remark 2.14. The previous theorem is only stated for connected trees, however, the same result with the same proof is true for forests. The B-function, in that case, is the product of B-functions for each connected component. For example, for $k = 2$ and forests with two connected components, Theorem 2.11 gives the joint probability distribution for two edges. Recall that $\mathbb{S}_n^2 \times \mathbb{T}^2$ can be interpreted as the set of directed edges, and note that the B-function of each edge is M . Then, the corresponding result says

$$\frac{1}{M^2} \sum |p_{a, (ij)(kl)} - \tilde{p}_{ij} \tilde{p}_{kl}| = o(1),$$

where the sum is over all disjoint pairs of directed edges, (i, j) and (k, l) , and

$$(2.9) \quad p_{a, (ij)(kl)} := E(\mathbf{1}_{(i, j) \in \mathcal{E}(\mathcal{G}_{\mathbf{a}})} (\mathcal{G}_{\mathbf{a}}) \mathbf{1}_{(k, l) \in \mathcal{E}(\mathcal{G}_{\mathbf{a}})} (\mathcal{G}_{\mathbf{a}})),$$

for $\mathcal{G}_{\mathbf{a}} = \mathcal{G}_{\mathbf{a}}(a, D)$.

Remark 2.15. Let us explore the weights now. Suppose that all $\tilde{p}_{ij} \simeq p$ are of the same order. Then it is not hard to see that $M \cdot \psi(s, T, D)$ is of order $(n^2 \cdot p) \cdot (n \cdot p)^{k-2} \simeq n^k \tilde{\mathbf{p}}(s, T)$, where $(s, T) \in \mathbb{S}^k \times \mathbb{T}^k$. Thus, $L_{\mathbf{a}}$ becomes

$$L_{\mathbf{a}}(a, k, D) \simeq \frac{1}{n^k} \sum_{(s, T) \in \mathbb{S}^k \times \mathbb{T}^k} \left| \frac{\mathbf{p}_{\mathbf{a}}(s, T)}{\tilde{\mathbf{p}}(s, T)} - 1 \right|.$$

In particular, if all degrees are d ($d_i = d$, for every $i \in [n]$ and $1 \leq d \leq n-1$), then symmetry implies that $p = \frac{d}{n-1}$, and that the independent model $\tilde{\mathcal{G}}(D)$ is Erdős–Rényi random graph $G(n, \frac{d}{n-1})$. The main point here is that the B- function gives the correct normalization.

Let us explore the statement of Theorem 2.11 for some small values of k . For $k = 2$,

$$(2.10) \quad L_{\mathbf{a}}(a, 2, D) = \frac{1}{M} \sum_{1 \leq i < j \leq n} |p_{a,(ij)} - \tilde{p}_{ij}| \leq c_k \cdot n^{\frac{-\nu}{2}+a} = o(1),$$

where $p_{a,(ij)} := E(\mathbf{1}_{\langle i,j \rangle \in \mathcal{E}(\mathcal{G}_{\mathbf{a}})}(\mathcal{G}_{\mathbf{a}}))$, and $\mathcal{G}_{\mathbf{a}} = \mathcal{G}_{\mathbf{a}}(a, D)$. Thus, the maximum entropy gives the edge probabilities in the random graph \mathcal{G}_a , i.e. $p_{a,(ij)} \sim \tilde{p}_{ij}$.

For $k = 3$, an ordered tree (s, T) in $\mathbb{S}_n^3 \times \mathbb{T}^3$ is a path of two edges, and its B- function is d_j , where j is the middle vertex of the tree $s(T)$. Then, by Remark 2.12,

$$(2.11) \quad \frac{L_{\mathbf{a}}(a, 3, n)}{3!} = \sum_{j \in [n]} \sum_{\{i,k\} \subset [n]} \frac{1}{M d_j} |p_{a,(ij)(kl)} - \tilde{p}_{ij} \tilde{p}_{jk}| = o(1),$$

where $p_{a,(ij)(kl)}$ is defined in Eq. (2.9). Next, we make an observation about Eq. (2.7).

Theorem 2.16. *Suppose that D is a strict graphic sequence of type ε as in Definition 2.3. Then the sum of variables in Theorem 2.11 is nearly constant:*

(1)

$$\frac{1}{M} \sum_{(s,T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} \mathbf{p}_{\mathbf{a}}(s, T) = k^{k-2} + O\left(\left(\frac{n}{M}\right)^{1-a}\right),$$

(2) and

$$\frac{1}{M} \sum_{(s,T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} \tilde{\mathbf{p}}(s, T) = k^{k-2} + O\left(\left(\frac{M}{n}\right)^{-\frac{1}{2}}\right),$$

(3) moreover,

$$(2.12) \quad \frac{1}{M} \sum_{(s,T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} (\mathbf{p}_{\mathbf{a}}(s, T) - \tilde{\mathbf{p}}(s, T)) = O\left(\left(\frac{n}{M}\right)^{1-a}\right).$$

Remark 2.17. The first part of Theorem 2.16 justifies the summation in the statement of Theorem 2.11, and shows that the weights do not overkill the summation. Moreover, Eq. (2.12) is Eq. (2.7) in Theorem 2.11 without the absolute sign. We will see later that the proof of the former equation is easier than the latter.

2.4. Graphs with a given degree sequence. Recall that \mathbb{G}^D is the set of all graphs with the degree sequence D . Let $\mathcal{G}_{\mathbf{g}}(D)$ be a random graph that is chosen uniformly out of the set \mathbb{G}^D . The random graph $\mathcal{G}_{\mathbf{g}}(D)$ is finer than the random graph $\mathcal{G}_{\mathbf{a}}(a, D)$ in the previous section. Because of a technical problem, we cannot provide the same result as in Theorem 2.11 in full generality. Therefore, we start with two conjectures. Then we continue with showing that the conjecture holds in

various regimes, such as very sparse graphs, dense graphs, and bipartite graphs. In addition, we see the relevance of the existing results in the literature, in each case.

Conjecture 2.18. *Suppose that the vector D is a strict graphic sequence of type ε (2.3), and that the entropy function $H_1(x)$ takes its maximum at $\tilde{\mathbf{p}}$, where $H_1(x)$ is defined in Eq. (2.1). Then, there exists a number $\eta > 0$ independent of n such that,*

$$(2.13) \quad e^{-\eta n \cdot \log(n)} \cdot e^{H_1(\tilde{\mathbf{p}})} \leq |\mathbb{G}^D| \leq e^{H_1(\tilde{\mathbf{p}})}.$$

We will see later that $P\left(\tilde{\mathcal{G}}(D) \in \mathbb{G}^D\right) = \frac{|\mathbb{G}^D|}{e^{H_1(\tilde{\mathbf{p}})}}$. So, Eq. (2.13) reads

$$(2.14) \quad e^{-\eta n \cdot \log(n)} \leq P\left(\tilde{\mathcal{G}}(D) \in \mathbb{G}^D\right) = \frac{|\mathbb{G}^D|}{e^{H_1(\tilde{\mathbf{p}})}} < 1.$$

This proves the upper bound that is the easy part of the Conjecture 2.18. However, the lower bound that is crucial for our next Conjecture is open.

Conjecture 2.19. *Suppose that the vector D is a strict graphic sequence of type ε . Let*

$$\mathbf{p}_{\mathbf{g}}(s, T) = E[\mathbf{1}_s(T, \mathcal{G}_{\mathbf{g}}(D))],$$

where $\mathbf{1}_s(T, G)$ is defined in Eq. 2.10. Define,

$$(2.15) \quad L_{\mathbf{g}}(k, D) := \frac{1}{M} \sum_{(s, T) \in \mathbb{S}_n^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} |\mathbf{p}_{\mathbf{g}}(s, T) - \tilde{\mathbf{p}}(s, T)|,$$

where \mathbb{T}^k , \mathbb{S}_n^k and $\psi(s, T, D)$ are defined in Section 2.2, and $\tilde{\mathbf{p}}(s, T)$ is as in Definition 2.10. Then we have,

$$L_{\mathbf{g}}(k, D) \leq c_k \cdot \sqrt{\frac{n \log(n)}{M}}.$$

Remark 2.20. Conjecture 2.18 implies Conjecture 2.19, i.e.

$$e^{-\eta n \cdot \log(n)} \leq P\left(\tilde{\mathcal{G}}(D) \in \mathbb{G}^D\right).$$

Thus, we have,

$$L_{\mathbf{g}}(k, D) \leq c_k \cdot \sqrt{\frac{n \log(n)}{M}}.$$

We will prove this in the Section 3.3.

2.5. Dense graphs with a given degree sequence: In this section, we present why Theorem 2.11 and Conjecture 2.19 are true for dense graphs.

Definition 2.21. For a positive number c_3 and numbers $c_1, c_2 \in (0, 1)$, we say that a graph with degree sequence D satisfies the *dense Erdős - Gallai conditions of type (c_1, c_2, c_3)* , if

(1) for every i

$$c_2(n-1) \leq d_i \leq c_1(n-1),$$

(2) and

$$\frac{1}{n^2} \inf_{B \subseteq \{1, \dots, n\}, |B| \geq c_2 n} \left\{ \sum_{j \notin B} \min\{d_j, |B|\} + |B|(|B| - 1) - \sum_{i \in B} d_i \right\} \geq c_3.$$

Remark 2.22. This definition implies that the vector D is δ -tame in the sense that it was defined in [1], which means $\delta \leq \tilde{p}_{ij} \leq 1 - \delta$ for $\delta(c_1, c_2, c_3)$, where $\tilde{\mathbf{p}} = (\tilde{p}_{ij})_{i,j}$ is the maximum entropy.

The next theorem is equivalent to the computations that appear in the last lines of the proof on page 34 of Theorem 1.1 in [6].

Theorem 2.23. *If the degree sequence D satisfies the dense Erdős-Rényi condition (the preceding definition), then the Conjectures 2.18 and 2.19 hold. Therefore,*

$$(2.16) \quad L_{\mathbf{g}}(k, D) \leq c_k \cdot (n \log(n))^{\frac{-1}{2}},$$

where $L_{\mathbf{g}}(k, D)$ is defined in Eq. (2.15). Moreover, for $\frac{1}{2} < a < \frac{1}{2} + \frac{1}{12}$, the Theorem 2.11 also holds, i.e.

$$(2.17) \quad L_{\mathbf{a}}(a, k, D) \leq c_k \cdot n^{\frac{-3}{4} + a},$$

where $L_{\mathbf{a}}(a, k, D)$ is defined in Eq. (2.7).

2.6. Very sparse graphs with a given degree sequence: Let us start this section with the very-sparseness definition.

Definition 2.24. A graph with a degree sequence D is called very sparse if $d_n^2 = d_{\max}^2 = o(M)$.

Now, we see the corresponding results to Theorem 2.11 and Conjecture 2.19 for very sparse random graphs. In addition, we see how our method is related to McKay's result.

Let us introduce new variables,

$$q_{ij} := \frac{d_i d_j}{M + d_i d_j}, \quad \text{where } 1 \leq i, j \leq n \text{ and } i \neq j.$$

Define $\mathcal{G}_{\mathbf{q}}(D)$ as a random graph with independent Bernoulli random edges with parameters q_{ij} . Let

$$(2.18) \quad \mathbf{p}_{\mathbf{q}}(s, T) = E[\mathbf{1}_s(T, \mathcal{G}_{\mathbf{q}}(D))],$$

where $\mathbf{1}_s(T, G)$ is defined in Eq. 2.10. This random graph is an alternative to the random graph $\tilde{\mathcal{G}}(D)$, which was constructed according to maximum entropy $\tilde{\mathbf{p}} = (\tilde{p}_{ij})_{i,j}$. The goal is to show that $\mathbf{p}_{\mathbf{q}}(s, T) \simeq \tilde{\mathbf{p}}(s, T)$, and in particular, $\tilde{p}_{ij} \simeq q_{ij}$.

Theorem 2.25. *Suppose that $d_n^2 = o(M)$, and that the vector D is a strict graphic sequence of type ε . Let $L_{\mathbf{a}}(a, k, D)$ and $L_{\mathbf{g}}(k, D)$ be as in Eq. (2.7) and (2.15), and also define*

$$(2.19) \quad L_{\mathbf{q}}(k, D) := \frac{1}{M} \sum_{(s, T) \in \mathbb{S}_n^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} |\mathbf{p}_{\mathbf{q}}(s, T) - \tilde{\mathbf{p}}(s, T)|,$$

where we used the notations in Theorem 2.11, and $\mathbf{p}_{\mathbf{q}}(s, T)$ is defined in Eq. (2.18). Then,

(1) for random graph $\mathcal{G}_{\mathbf{a}}$, we have

$$L_{\mathbf{a}}(a, k, D) \leq c_k \cdot \left(\frac{n}{M}\right)^{1/4} n^{a_1},$$

where $a_1 := a - \frac{1}{2}$ is such that $\left(\frac{n}{M}\right)^{1/2} n^{3a_1} \ll 1$.

(2) for random graph $\mathcal{G}_{\mathbf{g}}(D)$,

$$L_{\mathbf{g}}(k, D) \leq c_k \cdot \left(\left(\frac{n}{M}\right)^{1/4} n^{a_1} + \frac{d_n^2}{M}\right).$$

(3) moreover,

$$L_{\mathbf{q}}(k, D) \leq c_k \cdot \left(\left(\frac{n}{M}\right)^{1/4} n^{a_1} + \frac{d_n^2}{M}\right).$$

Remark. The enumeration method used in Gao et al. [18] compares the random graphs $\mathcal{G}_{\mathbf{g}}(D)$, and \mathcal{G}_q . Whereas, through out this paper, we place the random graph $\mathcal{G}_{\mathbf{g}}(D)$ in comparison with the random graph $\tilde{\mathcal{G}}(D)$, which comes from the maximum entropy. The previous theorem shows the relevance of the combinator method with respect to the general form by showing that $\tilde{p}_{ij} \simeq q_{ij}$ (that is when $k = 2$).

Remark. Note that $\sum_{j=1}^n q_{ij} \neq d_i$, so the point $q = (q_{ij})$ is not on the polytope \mathbb{P}^D . However, the condition $d_n^2 = o(M)$ demonstrates that this point is close to this polytope. In addition, the proof of the theorem depends on the combinatorial result of [27], so the condition $d_n^2 = o(M)$ is necessary.

2.7. Bipartite graphs with an exact degree sequence: In this section, we deal with bipartite graphs with a given degree sequence. Bipartite means that our graph has two vertex sets, namely set 1 and 2, and that there are no edges in between any two vertices of the same set. The advantage of the bipartite graphs is that Conjecture 2.18 holds, which is the subject of a paper of Barvinok [1]. They studied the number of 0-1 matrices with given row and column sums that are related to the problem of asymptotic enumeration for bipartite graphs, and the correspondence is given by the adjacency matrix of the graphs. This provides a way to state our results with no extra assumptions.

Let us start with a convention. We denote the parameters corresponding to each set of vertices by a number, either 1 or 2, in the subindex. Now, consider two integer vectors of $D_i = (d_{i,1}, \dots, d_{i,n_i})$, for $i = 1, 2$, and $d_{i,1} \leq \dots \leq d_{i,n_i}$. We say a bipartite graph has degree sequence (D_1, D_2) , if the degree sequence of the first vertex set is D_1 and the degree sequence of the second set is D_2 . Let $n := n_1 + n_2$ be the total number of vertices, and denote the set of all bipartite graphs with a given degree sequence (D_1, D_2) by \mathbb{G}^{D_1, D_2} . The condition $\sum_i d_{1,i} = \sum_j d_{2,j}$ is enough to assure that \mathbb{G}^{D_1, D_2} is non-empty.

Since the nature of a bipartite graph does not allow it to have some of the edges or some of the trees as its subgraph, we should change our entropy function slightly and restrict our set of connected trees. Therefore, while the setup for bipartite graphs is a little bit different from previous cases, the ideas are the same. We state some definitions.

Definition 2.26. Let us denote by $\mathcal{T}_{\mathbf{b}}(k, n)$ the ordered trees (s, T) in $\mathbb{S}_n^k \times \mathbb{T}^k$ such that $s(T)$ does not have any edges with both ends in either set 1 or set 2.

Definition 2.27. We let $\mathcal{G}_{\mathbf{b}}(D_1, D_2)$ be the random bipartite graph with a given degree sequence (D_1, D_2) that is uniformly chosen from the set \mathbb{G}^{D_1, D_2} . Also, for $(s, T) \in \mathcal{T}_{\mathbf{b}}^k$, let

$$(2.20) \quad \mathbf{p}_{\mathbf{b}}(s, T) = E[\mathbf{1}_s(T, \mathcal{G}_{\mathbf{b}}(D_1, D_2))],$$

where $\mathbf{1}_s(T, G)$ is defined in 2.10. We do not show the dependency of $\mathbf{p}_{\mathbf{b}}$ on D_1 and D_2 since it is understood.

For the maximum entropy we have the following new definition.

Definition 2.28. Let us consider the polytope \mathbb{P}^{D_1, D_2} of matrices $X = (x_{ij})$ such that $\sum_{j=1}^{n_2} x_{ij} = d_{i,1}$ for $1 \leq i \leq n_1$, and $\sum_{i=1}^{n_1} x_{ij} = d_{2,j}$ for $1 \leq j \leq n_2$, and $0 \leq x_{ij} \leq 1$ for all i, j . Also, define the entropy function for matrix $X = (x_{ij})$ as

$$H_2(x) = \sum_{i,j} H(x_{ij}) \text{ where } H(x) = -x \ln(x) - (1-x) \ln(1-x).$$

The entropy function $H_2(x)$ is strictly convex, hence, if the polytope \mathbb{P}^{D_1, D_2} has a non-empty interior, then the function $H_2(x)$ takes its maximum at some point $\tilde{\mathbf{p}} \in \mathbb{P}^{D_1, D_2}$.

Definition 2.29. Consider the maximum entropy $\tilde{\mathbf{p}} = (\tilde{p}_{ij})_{i,j}$, where $1 \leq i \leq n_1$, and $1 \leq j \leq n_2$. This time, we define the random graph $\tilde{\mathcal{G}}(D_1, D_2)$ as a random graph that has independent Bernoulli random edges with probability \tilde{p}_{ij} s, and let

$$(2.21) \quad \tilde{\mathbf{p}}(s, T) = E[\mathbf{1}_s(T, \tilde{\mathcal{G}}(D_1, D_2))].$$

Note that for the random graph $\tilde{\mathcal{G}}(D_1, D_2)$, there are no edges between the two vertices of set 1 or two vertices of set 2.

Theorem 2.30. Suppose that the polytope \mathbb{P}^{D_1, D_2} has a non-empty interior, and define

$$(2.22) \quad L_{\mathbf{b}}(k, D_1, D_2) := \frac{1}{M} \sum_{(s,T) \in \mathcal{T}_{\mathbf{b}}(k,n)} \frac{1}{\psi(s, T, D)} |\mathbf{p}_{\mathbf{b}}(s, T) - \tilde{\mathbf{p}}(s, T)|,$$

where we used the notations in Theorem 2.11, and $\mathbf{p}_{\mathbf{b}}(s, T)$ is defined in Eq. (2.20), and $D = (D_1, D_2)$. Then, we obtain

$$L_{\mathbf{b}}(k, D_1, D_2) \leq c_k \cdot \left(\frac{n \log(n)}{M} \right)^{\frac{1}{2}}.$$

Remark 2.31. Note that we are able to prove the above result in almost the full generality, whereas in Conjecture 2.19, we needed an extra lower bound, which is discussed in Remark 2.20. Moreover, it is also possible to formulate the counter part of Theorem 2.11 for bipartite graphs and prove it without any extra condition. However, that is a repetition of the previous work and so we skip it.

Remark 2.32. We observe that strict Erdős- Gallai conditions reduce to a much simpler equations in the case of bipartite graphs. However, dealing with that technicality is out of the scope of this paper, and we simply assume that \mathbb{P}^{D_1, D_2} has a non-empty interior.

2.8. Open questions. Here we list a couple of questions.

Question 1.: Can we prove our results under fewer assumptions? In particular, is it possible to drop the last two parts of Definition 2.3, and only use the Erdős - Gallai conditions for our theorems? (see Conjecture 3.13)

Question 2.: What can be said about triangle-counting in the sparse random graph G with a given degree sequence D ? How about other subgraphs?

Question 3.: What is the correct way of counting the subgraphs? In the sense that, can we define a metric or a topology for the space of sparse graphs using the weighted subgraph counts?

Question 4.: If the answer to the previous question is yes, are there any limiting objects under that metric?

Question 5.: Is there any way to define a limiting object for graphs with a given degree sequence using the maximum entropy?

3. PROOFS

The rest of the paper is organized as follows. In Section 3.1, we prove Theorems 2.16 and Theorem 2.11. For that, we use Theorem A.1 that is presented at the appendix with its proof. In Section 3.3, we see the proof of Remark 2.20. Then we prove Theorems 2.23, 2.25, and 2.30, in Sections 3.4, 3.5, and 3.6, respectively. We also use the notations in Sections 2.1 and 2.2 frequently.

3.1. Graph with an almost given degree sequence. The aim of this section is to present the proof of Theorem 2.11, which will be completed in Subsection 3.1.2. Theorem 2.16 is required for this goal, so we start with a proof of that.

3.1.1. Proof of Theorem 2.16: We see a few lemmas, and then the proof of Theorem 2.16.

Lemma 3.1. (Upper bound). *Let $T \in \mathbb{T}^k$ and $G \in \mathbb{G}^D$. If*

$$F(T, G) = \sum_s \frac{1}{\psi(s, T, G)} \prod_{e=\langle u_1, u_2 \rangle \in \mathcal{E}(T)} \mathbf{1}_{\langle s(u_1), s(u_2) \rangle \in \mathcal{E}(G)}(G),$$

then $F(T, G) \leq M$.

Proof. We prove it by induction on the size k of the tree. We show that for any tree T with k vertices there is a tree T' with one vertex less such that

$$F(T, G) \leq F(T', G),$$

for all G . Take any leaf u_1 of the tree, and suppose that u_1 is linked to u_2 by an edge. Note that $b_{u_1} = 1$. Let T_{k-1} be the tree obtained by deleting u_1 and the linking edge. If we have already chosen $s(u)$ for $u \in T_{k-1}$, then we can first do the summation

$$\sum_x \mathbf{1}_{\langle s(u_2), x \rangle \in \mathcal{E}(G)}(G) \leq d_{s(u_2)}.$$

The degrees for the vertices in the new tree T' are all the same as in T except for u_2 , and b_{u_2} is reduced by 1. Therefore

$$F(T_k, G) \leq F(T_{k-1}, G),$$

and $F(T_2, G) = M$. \square

Lemma 3.2. (Lower bound). *For F , T and G , as in the previous lemma,*

$$F(T, G) \geq M - \frac{nk(k-1)}{2}.$$

Proof. Let v be any leaf of T , and let u be such that $\langle u, v \rangle$ is the only edge of v . Since at most $k-1$ possible edges from $s(u)$ could lead to vertices of $\mathcal{V}(s(T)) \setminus \{s(v)\}$, we write the inequality

$$\sum_{\{x: x \neq s(v'), \forall v' \in \mathcal{V}(T), v' \neq v\}} \mathbf{1}_{\langle s(u), x \rangle \in \mathcal{E}(G)}(G) \geq d_{s(u)} - (k-1).$$

This provides a recurrence relation

$$\begin{aligned} & F(T_{k-1}, G) - F(T_k, G) \\ & \leq (k-1) \left[\sum_s \frac{1}{d_{s(u)}} \frac{1}{\psi(s, T_{k-1}, G)} \prod_{e=\langle u_1, u_2 \rangle \in \mathcal{E}(T_{k-1})} \mathbf{1}_{\langle s(u_1), s(u_2) \rangle \in \mathcal{E}(G)}(G) \right], \end{aligned}$$

where T_{k-1} is T_k when we remove v and the edge attached to it from T_k . Let us denote by $H(T_{k-1}, G)$ the right hand side of the above formula. Now, we can bound $H(T_{k-1}, G)$, as in the previous lemma. So, there is a tree T_{k-2} by removing a vertex of T_{k-1} such that

$$H(T_{k-2}, G) \leq H(T_{k-1}, G).$$

Continuing with that procedure, we can also assume without loss of generality that u is the last vertex to be removed, i.e. $T_1 = \{u\}$ in the sequence. Now the upper bound can be estimated and the last step is $\sum_i 1 = n$, rather than $\sum_i d_i = M$, because of the extra term $\frac{1}{d_{s(u)}}$ in H . Providing us with the estimate

$$F(T_{k-1}, G) - F(T_k, G) \leq (k-1)n,$$

Which yields

$$F(T, G) = F(T_k, G) \geq M - \frac{nk(k-1)}{2}.$$

\square

If we use the degree sequence $\{d_i\}$ in the definition of the B- function while the actual degrees are some what different $\{\tilde{d}_i\}$ that satisfies

$$|d_i - \tilde{d}_i| \leq cd_i^a,$$

i.e. $G \in \mathbb{G}_a^D$ then we need an error bound on the difference

$$(3.1) \quad Z_k = \sum_s \left| \prod_{u \in \mathcal{V}(T)} \frac{1}{[d_{s(u)}]^{b_u-1}} - \prod_{u \in \mathcal{V}(T)} \frac{1}{[\tilde{d}_{s(u)}]^{b_u-1}} \right| \mathbf{1}_s(T, G),$$

where

$$\mathbf{1}_s(T, G) = \prod_{\langle u_1, u_2 \rangle \in \mathcal{E}(T)} \mathbf{1}_{\langle s(u_1), s(u_2) \rangle \in \mathcal{E}(G)}(G).$$

Lemma 3.3. *For Z_k as in Eq. (3.1), we have*

$$Z_k \leq c k M^a n^{1-a}.$$

Proof. Summing over choices of $s(u_1)$ where u_1 is a leaf connected through the vertex u_2 ,

$$\begin{aligned} Z_k &\leq \sum_s \left| \prod_{u \in \mathcal{V}(T_k)} \frac{1}{[d_s(u)]^{b_u-1}} - \prod_{u \in \mathcal{V}(T_k)} \frac{1}{[\tilde{d}_s(u)]^{b_u-1}} \right| \mathbf{1}_s(T_k, G) \\ &\leq \sum_s \left| \frac{\tilde{d}_s(u_2)}{d_s(u_2)} \prod_{u \in \mathcal{V}(T_{k-1})} \frac{1}{[d_s(u)]^{b_u-1}} - \prod_{u \in \mathcal{V}(T_{k-1})} \frac{1}{[\tilde{d}_s(u)]^{b_u-1}} \right| \mathbf{1}_s(T_{k-1}, G) \\ &\leq Z_{k-1} + \sum_s \left| \frac{\tilde{d}_s(u_2)}{d_s(u_2)} - 1 \right| \prod_{u \in \mathcal{V}(T_{k-1})} \frac{1}{[d_s(u)]^{b_u-1}} \mathbf{1}_s(T_{k-1}, G) \\ &\leq Z_{k-1} + c \sum_s \prod_{u \in \mathcal{V}(T_{k-1})} \frac{d_s^{a-1}(u_2)}{[d_s(u)]^{b_u-1}} \mathbf{1}_s(T_{k-1}, G). \end{aligned}$$

Let us concentrate on the second term. We can assume with out loss of generality that $\tilde{d}_v \leq c d_v$ for all $v \in G$. Then, since $b(u) \geq 1$,

$$Z_k \leq Z_{k-1} + c \sum_s \prod_{u \in \mathcal{V}(T_{k-1})} \frac{\tilde{d}_s^{a-1}(u_2)}{[\tilde{d}_s(u)]^{b_u-1}} \mathbf{1}_s(T_{k-1}, G).$$

Successive summation ends up with $T_1 = \{u_2\}$. That leaves us with

$$Z_k \leq Z_{k-1} + c \sum_v \tilde{d}_v^a,$$

and

$$\sum_v \tilde{d}_v^a \leq c \sum_v d_v^a \leq c M^a n^{1-a}.$$

Summing up, we get

$$Z_k \leq c k M^a n^{1-a}.$$

□

Proof of Theorem 2.16. Using the previous three lemmas, and for the random graph \mathcal{G}_a with an almost given degree sequence D that is uniform over all graphs in \mathbb{G}_a^D , we write

$$\begin{aligned} F(T) &:= \frac{1}{M} \cdot E \left[\sum_s \prod_{u \in \mathcal{V}(T)} \frac{1}{[\tilde{d}_s(u)]^{b_u-1}} \mathbf{1}_s(T, \mathcal{G}_a) \right] \\ &= \frac{1}{M} \cdot E_a \left[\sum_s \prod_{u \in \mathcal{V}(T)} \frac{1}{[d_s(u)]^{b_u-1}} \mathbf{1}_s(T, \mathcal{G}_a) \right] \\ &= 1 + O\left(\frac{n}{M}\right) + O\left(\left(\frac{n}{M}\right)^{1-a}\right), \end{aligned}$$

where the constant in the O notation depends on k as n goes to infinity. Recall from Cayley's theorem that \mathbb{T}^k has $k^{(k-2)}$ elements. Thus,

$$\frac{1}{M} \sum_{(s,T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s,T,D)} \mathbf{p}_a(s,T) = k^{(k-2)} + O\left(\frac{n}{M}\right) + O\left(\left(\frac{n}{M}\right)^{1-a}\right).$$

We notice that k is constant as n goes to infinity, and that $\frac{1}{2} \leq a < 1$.

For the second part and for the random graph $\tilde{\mathcal{G}}(D)$ with independent Bernoulli edges and parameters (\tilde{p}_{ij}) , we observe that

$$\begin{aligned} E \left[\mathbf{1}_s(T, \tilde{\mathcal{G}}(D)) \right] &= \prod_{\langle u_1, u_2 \rangle \in \mathcal{E}(T)} E \left[\mathbf{1}_{\langle s(u_1), s(u_2) \rangle \in \mathcal{E}(\tilde{\mathcal{G}}(D))}(\tilde{\mathcal{G}}(D)) \right] \\ &= \prod_{\langle u_1, u_2 \rangle \in \mathcal{E}(T)} \tilde{p}_{\langle s(u_1), s(u_2) \rangle}. \end{aligned}$$

In addition, $d_i = \sum_{j \in [n] \setminus \{i\}} \tilde{p}_{\langle i, j \rangle}$. Incorporating these two estimates in the proof from previous part we get the second part, and third part of the theorem is the combination of the first two parts and $a > \frac{1}{2}$. \square

Now, we are ready to prove the only remaining result of Section 2.3.

3.1.2. Proof of Theorem 2.11: First, we prove Proposition 2.1 and a series of Lemmas that are all required for the proof of Theorem 2.11. Lemma 3.8 is important because it is related to the technical condition given in the theorem. In addition, we prove Theorem 2.11 under weaker assumptions that are conjectured to hold. Furthermore, here and in the pages that follow, c plays the role of a general constant, and c_k is a constant that depends on k .

Proof of Proposition 2.1. Suppose that the vector D satisfies the strict Erdős-Gallai condition, then lemma (12.2) in [1] implies that the polytope \mathbb{P}^D has a non-empty interior. The reverse direction has two steps, and we assume that \mathbb{P}^D has a non-empty interior. First, an argument of Lagrange multipliers shows a relation between parameters \tilde{p}_{ij} s (for details look at the beginning of the proof of Theorem 2.1, [1]). It is easy to see that the entropy function $H_1(x)$ is strictly concave. The polytope \mathbb{P}^D is compact and H_1 attains its unique maximum $\tilde{p} = (\tilde{p}_{ij})$. Moreover, \tilde{p} is in the interior of \mathbb{P}^D , $0 < \tilde{p}_{ij} < 1$. Therefore, the gradient of H_1 should be perpendicular to \mathbb{P}^D , so there is a vector $\vec{\lambda} \in \mathbb{R}^n$ such that,

$$(3.2) \quad \partial_{ij} H_1 = \log \left(\frac{1 - x_{ij}}{x_{ij}} \right) = \lambda_i + \lambda_j.$$

If we let

$$(3.3) \quad r_i = e^{-\lambda_i},$$

and rewrite (3.2) in terms of r_i s, then we get for every i and j ,

$$(3.4) \quad \tilde{p}_{ij} = \frac{r_i r_j}{1 + r_i r_j}.$$

The point \tilde{p} is a point in \mathbb{P}^D , so for every $i \in [n]$,

$$d_i = \sum_{k \neq i} \frac{r_i r_k}{1 + r_i r_k}.$$

Second, we use the above relation to get (2.2). We expand the left hand side of (2.2) in terms of \tilde{p}_{ij} and get

$$\begin{aligned} \sum_{i=n-k+1}^n d_i &= \sum_{i=n-k+1}^n \sum_{l \neq i} \frac{r_i r_l}{1 + r_i r_l} \\ &= \left(\sum_{i=n-k+1}^n \sum_{l > n-k} + \sum_{i=n-k+1}^n \sum_{l \leq n-k} \right) \frac{r_i r_l}{1 + r_i r_l}. \end{aligned}$$

Now, the first sum has at most $k(k-1)$ terms and each term is strictly less than one. Again for any l in the second term, we have $\sum_{i=n-k+1}^n \frac{r_i r_l}{1 + r_i r_l} < k$, and $\sum_{i=n-k+1}^n \frac{r_i r_l}{1 + r_i r_l} \leq \sum_{i \neq l} \frac{r_i r_l}{1 + r_i r_l} = d_l$. Therefore, the sum is less than or equal to $\min\{k, d_l\}$, and

$$\sum_{i=n-k+1}^n d_i < k(k-1) + \sum_{l=1}^{n-k} \min\{k, d_l\}.$$

□

Recall from Section 2.1 that the random graph $\tilde{\mathcal{G}}(D)$ is a collection of independent Bernoulli random edges with parameter \tilde{p}_{ij} , and we just showed that $\tilde{p}_{ij} = \frac{r_i r_j}{1 + r_i r_j}$, where r_i are the same as in the past proof. Moreover, we remember that $\mathcal{G}_{\mathbf{a}}(a, D)$ is drawn uniformly out of \mathbb{G}_a^D . The following lemma and propositions deal with changing the underlying measure of $\mathcal{G}_{\mathbf{a}}(a, D)$ to that of $\tilde{\mathcal{G}}(D)$.

Proposition 3.4. *Let us suppose that \tilde{p}_{ij} s satisfy $\delta \leq \tilde{p}_{ij} \leq 1 - \delta$, for $i, j \in [n]$, and some $\delta > 0$. Then, for large values of n , we have*

$$P\left(\tilde{\mathcal{G}}(D) \in \mathbb{G}_a^D\right) > e^{-10 \cdot \mathcal{C}_1(D)},$$

where,

$$(3.5) \quad \mathcal{C}_1(D) := |\log(\delta)| \cdot n \cdot \log^{\frac{10}{a_1}}(n),$$

and $a_1 := a - \frac{1}{2}$.

Proof. The proof is long, and we leave it for the next section. □

Lemma 3.5. *For a graph G with the degree sequence $D(G) = (d_1(G), \dots, d_n(G))$, we have*

$$P\left(\tilde{\mathcal{G}}(D) = G\right) = \frac{\prod_{i=1}^n r_i^{d_i(G)}}{\prod_{i < j} (1 + r_i r_j)}.$$

Proof. Let V and E be the vertex and edge sets of G , then

$$\begin{aligned} (3.6) \quad P\left(\tilde{\mathcal{G}}(D) = G\right) &= \prod_{\langle i, j \rangle \in E} \tilde{p}_{ij} \prod_{\langle i, j \rangle \notin E} (1 - \tilde{p}_{ij}) \\ &= \prod_{\langle i, j \rangle \in E} \frac{r_i r_j}{1 + r_i r_j} \prod_{\langle i, j \rangle \notin E} \frac{1}{1 + r_i r_j} \\ &= \frac{\prod_{\langle i, j \rangle \in E} r_i r_j}{\prod_{i < j} (1 + r_i r_j)}. \end{aligned}$$

But, for every $1 \leq i \leq n$, the number of pairs $\langle i, j \rangle$ in E is exactly the degree of vertex i in G that is equal to $d_i(G)$. Therefore, the numerator of Eq. (3.6) becomes $\prod_{1 \leq i \leq n} r_i^{d_i(G)}$. \square

Proposition 3.6. *For any subset A of \mathbb{G}_a^D , we have,*

$$P(\mathcal{G}_a(a, D) \in A) \leq e^{2\mathcal{C}_2(D)} \frac{P(\tilde{\mathcal{G}}(D) \in A)}{P(\tilde{\mathcal{G}}(D) \in \mathbb{G}_a^D)},$$

where

$$(3.7) \quad \mathcal{C}_2(D) := \sum_{i=1}^n d_i^a |\log(r_i)|.$$

Proof. Define $\tilde{P}_{\min} := \min_{G \in \mathbb{G}_a^D} P(\tilde{\mathcal{G}} = G)$ and $\tilde{P}_{\max} := \max_{G \in \mathbb{G}_a^D} P(\tilde{\mathcal{G}} = G)$, and let \tilde{P}_{\min} and \tilde{P}_{\max} be achieved for graphs G_1 and G_2 with degree sequences $D(G_1)$ and $D(G_2)$. In addition, because $D(G_1)$ and $D(G_2)$ are in \mathbb{G}_a^D ,

$$|d_i(G_1) - d_i(G_2)| \leq |d_i(G_1) - d_i| + |d_i - d_i(G_2)| \leq 2d_i^a,$$

and based on the previous lemma,

$$(3.8) \quad \begin{aligned} \left| \log\left(\frac{\tilde{P}_{\max}}{\tilde{P}_{\min}}\right) \right| &= \sum_{i=1}^n |(d_i(G_1) - d_i(G_2)) \log(r_i)| \\ &\leq 2 \sum_{i=1}^n d_i^a |\log(r_i)| = 2\mathcal{C}_2(D). \end{aligned}$$

Now, for any graph $G \in A$, we have

$$P(\mathcal{G}_a(a, D) = G) = \frac{1}{|\mathbb{G}_a^D|} \leq \frac{1}{\tilde{P}_{\min}} \frac{P(\tilde{\mathcal{G}}(D) = G)}{|\mathbb{G}_a^D|}.$$

Furthermore, $P(\tilde{\mathcal{G}}(D) \in \mathbb{G}_a^D) \leq \tilde{P}_{\max} \cdot |\mathbb{G}_a^D|$, so

$$P(\mathcal{G}_a(a, D) \in A) \leq \frac{\tilde{P}_{\max}}{\tilde{P}_{\min}} \frac{1}{P(\tilde{\mathcal{G}} \in \mathbb{G}_a^D)} P(\tilde{\mathcal{G}} \in A) \leq \frac{e^{2\mathcal{C}_2(D)}}{P(\tilde{\mathcal{G}} \in \mathbb{G}_a^D)} P(\tilde{\mathcal{G}} \in A),$$

where $\tilde{\mathcal{G}} = \tilde{\mathcal{G}}(D)$. \square

We continue with a lemma that investigates some properties of r_i s that are required frequently through out this section.

Lemma 3.7. *We suppose that $d_1 \leq \dots \leq d_n$, then,*

- a) $r_1 \leq \dots \leq r_n$,
- b) and $r_1 r_n > \frac{1}{n}$.
- c) If $r_k \geq 1$, for some $1 \leq k \leq n$, then $r_{k+1}/r_k < n^4$.
- d) If $r_k > n^2$, for some $1 \leq k \leq n$, then $\sum d_i \leq \frac{1}{2}M$, where the sum is over $1 \leq i \leq n - d_k - 1$.

Proof. We leave the proof for Appendix B. \square

Lemma 3.8. *Suppose that the vector D is a strict graphic sequence of type (ε, ν) as in Definition 2.7. Then,*

$$(3.9) \quad \max \{ \mathcal{C}_1(D), \mathcal{C}_2(D) \} \leq 8 \cdot n^{a-\nu-\frac{1}{2}} \cdot M \cdot \log(n),$$

where $\mathcal{C}_1(D)$ and $\mathcal{C}_2(D)$ are defined in Eq. (3.5) and (3.7), respectively, and $\frac{1}{2} < a < \frac{\nu}{4} + \frac{1}{2}$.

Proof. Let $A_k := \{n - d_k, \dots, n\}$, and recall that

$$\ell(D) := \max \left\{ k \in [n] \mid \sum_{i \in A_k} d_i \leq \frac{M}{2} \right\},$$

as in Eq. (2.5). Correspondingly, if we denote $\ell(D)$ by ℓ , then we have $\sum_{i \in A_\ell} d_i \leq \frac{M}{2}$, and $\sum_{i \notin A_\ell} d_i \geq \frac{M}{2}$. Now, part d of Lemma 3.7 implies that $r_\ell \leq n^2$. By part a of the same lemma, we get $d_i = d_{i+1}$, whenever $r_i = r_{i+1}$. Therefore, the set $I := \{r_i \mid i \geq \ell\}$ of distinct values among r_ℓ, \dots, r_n can be at most the distinct values among d_ℓ, \dots, d_n and has at most $d_n - d_\ell + 1$ elements. Again, by part c of that lemma, each r_i is either less than 1 or can exceed the previous one by at most n^4 , hence

$$(3.10) \quad \log(r_n) \leq \log(n) \cdot (2 + 4(d_n - d_\ell)).$$

Finally, part b of that lemma gives a lower bound for r_1 that is

$$(3.11) \quad -\log(r_1) \leq \log(n) \cdot (3 + 4(d_n - d_\ell)).$$

Combining the previous equations and part a of Lemma 3.7, we have

$$(3.12) \quad |\log(r_i)| < 4 \cdot \log(n) \cdot (d_n - d_\ell + 1).$$

We note that $\tilde{p}_{ij} = \frac{r_i r_j}{1 + r_i r_j}$, for $1 \leq i < j \leq n$, and we write

$$\delta_{ij} := \min \left\{ \frac{r_i r_j}{1 + r_i r_j}, \frac{1}{1 + r_i r_j} \right\}.$$

Then, $|\log(\delta_{ij})| \leq \log(2) + |\log(r_j)| + |\log(r_i)|$. Since r_i s are increasing (Lemma 3.7), and by Eq. (3.10) and (3.11), we have $|\log(\delta_{ij})| \leq 8 \log(n) \cdot (d_n - d_\ell + 1)$.

According to Assumption 2.7, $(d_n - d_\ell + 1)$ is bounded by $\sqrt{\frac{M}{n}} n^{-\nu}$, and $M \geq n^{1+\varepsilon}$. Hence, for $\delta = \min_{i,j} \delta_{ij}$,

$$\begin{aligned} \mathcal{C}_1(D) &= |\log(\delta)| \cdot n \cdot \log^{\frac{10}{a}}(n) \\ &\leq 8 \sqrt{\frac{M}{n}} n^{-\nu} \cdot n \cdot \log^{\frac{10}{a}+1}(n) \\ &\leq 8M \cdot n^{-\nu} \frac{\log^{\frac{10}{a}+1}(n)}{n^{\frac{\varepsilon}{2}}}. \end{aligned}$$

Therefore, for large enough n , $\mathcal{C}_1(D)$ is less than $8 \cdot M \cdot n^{-\nu}$.

Next, we apply Eq. (3.12) for $\mathcal{C}_2(D)$ to get

$$\mathcal{C}_2(D) = \sum_{i=1}^n d_i^a |\log(r_i)| < 4 \cdot \log(n) \cdot (d_n - d_\ell + 1) \sum_{i=1}^n d_i^a.$$

An application of Holder inequality gives the bound $n \cdot \left(\frac{M}{n}\right)^a$ for the term $\sum_{i=1}^n d_i^a$ in the above equation. Thus,

$$\begin{aligned}
 \mathcal{C}_2(D) &< 4 \cdot \log(n) \cdot (d_n - d_\ell + 1) \cdot n \cdot \left(\frac{M}{n}\right)^a \\
 (3.13) \quad &\leq 4 \cdot \log(n) \cdot n^{-\nu} \cdot \sqrt{\frac{M}{n}} \cdot n \cdot \left(\frac{M}{n}\right)^a \\
 &= 4 \cdot \log(n) \cdot n^{-\nu} \cdot M \cdot \left(\frac{M}{n}\right)^{a-\frac{1}{2}}
 \end{aligned}$$

Since $M \leq n^2$, we get $\mathcal{C}_2(D) = 4 \cdot n^{a-\nu-\frac{1}{2}} \cdot M \cdot \log(n)$. □

Proposition 3.9. *Suppose that A is a subset of $\mathbb{S}_n^k \times \mathbb{T}^k$, then*

$$P \left(\left| \sum_{(s,T) \in A} \frac{1}{\psi(s,T,D)} \left(\mathbf{1}_s(T, \tilde{\mathcal{G}}(D)) - \tilde{\mathbf{p}}_s(T, \tilde{\mathcal{G}}(D)) \right) \right| > \mu \epsilon \right) \leq e^{-(c\mu M)\epsilon^2}$$

where

$$\mu := \frac{1}{M} \sum_{(s,T) \in A} \frac{1}{\psi(s,T,D)} \tilde{\mathbf{p}}(s,T),$$

and again,

$$\tilde{\mathbf{p}}(s,T) := E \left[\mathbf{1}_s(T, \tilde{\mathcal{G}}(D)) \right],$$

and $M = \sum_{i=1}^n d_i$.

We borrow an idea from Janson's paper [21] to build a concentration result for the proof of the above lemma. However, the proof of our concentration inequality is rather long and is left for Appendix A. Next, we state some notations and the proof of Lemma 3.9.

Remember from the beginning of Section 3.1 that an ordered tree is a combination of an injective function and a tree, i.e. $(s,T) \in \mathbb{T}^k \times \mathbb{S}^k$, and the tree is understood as the image of T under s . We wish to define the union of such ordered trees.

Definition 3.10. We consider two ordered trees (s_1, T_1) and (s_2, T_2) of size k_1 and k_2 , respectively, whose edge sets intersect i.e. $\mathcal{E}(s_1(T_1)) \cap \mathcal{E}(s_2(T_2)) \neq \emptyset$. We define their wedge sum as follows: we let H be the union of edges of the graphs $s_1(T_1)$ and $s_2(T_2)$, with vertex set

$$V := \mathcal{V}(s_1(T_1)) \cup \mathcal{V}(s_2(T_2)) = \{w_1, \dots, w_{k_3}\} \subseteq \{1, \dots, n\},$$

for some integer k_3 , and we ordered w_i s according to their order in $\{1, \dots, n\}$. Then, we fix a common edge e of $s_1(T_1)$ and $s_2(T_2)$, i.e. an edge $e \in \mathcal{E}(s_1(T_1)) \cap \mathcal{E}(s_2(T_2))$. For any vertex $v \in \mathcal{V}(s_1(T_1)) \cap \mathcal{V}(s_2(T_2))$ that is not the end point of any edges of $\mathcal{E}(s_1(T_1)) \cap \mathcal{E}(s_2(T_2))$, there is a path P_v that attaches v to the edge e with edges in the tree $s_2(T_2)$. We let $\langle v, v_1 \rangle$ be the first edge in that path, and we erase it. Since $v \in \mathcal{V}(s_1(T_1))$, v is still connected to e through the edges of $s_1(T_1)$, and also, the vertex v_1 is connected to the edge e via the rest of the path P_v . Therefore, our resulting graph is connected.

For any remaining loop, we delete an edge of it that lays in $\mathcal{E}(s_2(T_2)) \setminus \mathcal{E}(s_1(T_1))$, and we continue this process until all loops are exhausted. This does not make our

graph disconnected, since we only erase one edge from a loop. Let H' be the reduced version of H .

Next, we define $s_3 : [k_3] \rightarrow [n]$ as $s(i) = w_i$, for $1 \leq i \leq k_3$, and we let T_3 be a connected tree with k_3 vertices such that $s_3(T_3)$ is H' . Then, (s_3, T_3) is the wedge sum of (s_1, T_1) and (s_2, T_2) , and we use the notation

$$(s_3, T_3) := (s_1, T_1) \vee (s_2, T_2).$$

Remark 3.11. Note that the wedge sum is not an injective function. Suppose we are given the wedge sum (s_3, T_3) , and we would like to retrieve the two ordered trees (s_1, T_1) and (s_2, T_2) . A simple but crude bound on the number of such pairs of ordered trees is $2^{k_3 \cdot (k_3 - 1)} \cdot (k_3!)^2$, after all, $H_1 = s_1(T_1)$, $H_2 = s_2(T_2)$ are two subgraphs with vertex sets in $\mathcal{V}(s_3(T_3))$, and knowing H_1 and H_2 , there are at most $k_3!$ to choose either of s_1 or s_2 .

We now look at the behavior of B -function under the wedge sum.

Lemma 3.12. *For any $G \in \mathbb{G}_n$,*

$$(3.14) \quad \psi(s_3, T_3, G) \leq \psi(s_1, T_1, G) \psi(s_2, T_2, G),$$

where $(s_3, T_3) = (s_1, T_1) \vee (s_2, T_2)$.

Proof. Recall that the B -function is

$$\psi(s, T, G) = \prod_{u \in \mathcal{V}(T)} d_{s(u)}^{b_u - 1},$$

where $D = (d_i)_{i \in [n]}$ is the degree sequence of G , and b_u is the degree of a vertex u in the graph T . In addition, for $w \in V := \mathcal{V}(s_3(T_3))$ and $i \in \{1, 2, 3\}$, we let $c_i(w)$ be the degree of $s_i^{-1}(w)$ in T_i if that exists, and 1 otherwise, i.e. $c_i(w) = \max\{b_{s_i^{-1}(w)}, 1\}$. Therefore, we can rewrite the B -function as

$$\psi(s_i, T_i, G) = \prod_{w \in V} d_w^{c_i(w) - 1}.$$

Now, we only need to show that $c_3(w) + 1 \leq c_1(w) + c_2(w)$. We observe that $s(T_3)$ is the subset of the union of edges of $s(T_1)$ and $s(T_2)$. So, the problem may only arise at a vertex w in $\mathcal{V}(s_1(T_1)) \cap \mathcal{V}(s_2(T_2))$ that is not the end point of any edges of $\mathcal{E}(s_1(T_1)) \cap \mathcal{E}(s_2(T_2))$. But we erased an edge from any of these vertices, due to our construction. That completes the proof. \square

Proof of Proposition 3.9. We are using Theorem A.1, which is proven at the appendix A, with parameters $\{J_i\} = \{\mathbf{1}_{\langle i, j \rangle}\}$, $Q = \mathbb{S}^k \times \mathbb{T}^k$, $\alpha = (s, T) \in A$, and $\omega_{(s, T)} = M \cdot \psi(s, T, D)$. In addition,

$$p_\alpha = E[\mathbf{1}_\alpha] = E\left[\mathbf{1}_s\left(T, \tilde{\mathcal{G}}(D)\right)\right] = \tilde{\mathbf{p}}(s, T),$$

and for $S = \sum_{\alpha \in A} \frac{1}{\omega_\alpha} \mathbf{1}_\alpha$,

$$(3.15) \quad \lambda = E[S] = \sum_{\alpha} \frac{1}{\omega_\alpha} p_\alpha = \frac{1}{M} \sum_{(s, T) \in A} \tilde{\mathbf{p}}(s, T) = \mu.$$

All we need is to show that δ_1 and δ_2 in the statement of Theorem A.1 are bounded by $\frac{c}{M}$, where c is a constant depending on k . For δ_1 we have,

$$\begin{aligned}
 (3.16) \quad \delta_1 &= \frac{1}{\lambda} \sum_{\alpha} \frac{p_{\alpha}}{\omega_{\alpha}^2} \\
 &= \frac{1}{\mu} \sum_{(s,T) \in A} \frac{1}{M^2 \psi^2(s,T,D)} \tilde{\mathbf{p}}(s,T) \\
 &\leq \frac{1}{M^2} \frac{1}{\mu} \sum_{(s,T) \in A} \frac{1}{\psi(s,T,D)} \tilde{\mathbf{p}}(s,T) \\
 &\leq \frac{1}{M}.
 \end{aligned}$$

For the last step, we used an argument similar to Lemma 3.1 to get

$$(3.17) \quad \frac{1}{\mu} \sum_{(s,T) \in A} \frac{1}{\psi(s,T,D)} \tilde{\mathbf{p}}(s,T) \leq M.$$

Computing δ_2 needs more work. Two ordered trees are dependent, $s_1(T_1) \sim s_2(T_2)$ ($\alpha \sim \beta$), iff they share at least one edge, or $\mathcal{E}(s_1(T_1)) \cap \mathcal{E}(s_2(T_2)) \neq \emptyset$. We let $\tilde{\mathcal{G}} = \tilde{\mathcal{G}}(D)$, and $N(s_1, T_1) := \{(s_2, T_2) \mid s_1(T_1) \sim s_2(T_2)\}$. It follows by Lemma 3.12 that

$$\begin{aligned}
 (3.18) \quad \delta_2 &= \frac{1}{\lambda} \sum_{\alpha} \sum_{\beta \sim \alpha} \frac{1}{\omega_{\alpha} \omega_{\beta}} E I_{\alpha} I_{\beta} \\
 &= \frac{1}{\mu M^2} \sum_{(s_1, T_1) \in A} \sum_{(s_2, T_2) \in N(s_1, T_1)} \frac{1}{\psi(s_1, T_1, D) \psi(s_2, T_2, D)} E \left[\mathbf{1}_{s_1}(T_1, \tilde{\mathcal{G}}) \mathbf{1}_{s_2}(T_2, \tilde{\mathcal{G}}) \right] \\
 &\leq \frac{1}{\mu \cdot M^2} \sum_{(s_1, T_1) \in A} \sum_{(s_2, T_2) \in N(s_1, T_1)} \frac{1}{\psi(s_3, T_3, D)} E \left[\mathbf{1}_{s_3}(T_3, \tilde{\mathcal{G}}) \right],
 \end{aligned}$$

where $(s_3, T_3) = (s_1, T_1) \vee (s_2, T_2)$ in the last sum, and we denote the set of all such ordered trees (s_3, T_3) by $U(s_1, T_1)$. Moreover, for $(s, T) \in \mathbb{S}^j \times \mathbb{T}^j$ and $j > k$, we let $T^{(k)}$ be the restriction of T to its first k vertices and we define

$$U(s_1, T_1, j) := \left\{ (s, T) \in \mathbb{S}^j \times \mathbb{T}^j \mid s(T^{(k)}) = s_1(T_1) \right\}.$$

In other words, $U(s_1, T_1, j)$ is the set of ordered trees of size j that are extensions of (s_1, T_1) .

Now, the wedge sum tree $(s_1, T_1) \vee (s_2, T_2)$ is in $U(s_1, T_1, j)$, for some $k \leq j \leq 2(k-1)$. Alternatively, for each element (s_3, T_3) of $U(s_1, T_1, j)$ and by Remark 3.11, the number of pairs (s_1, T_1) and (s_2, T_2) such that $(s_3, T_3) = (s_1, T_1) \vee (s_2, T_2)$ is bounded by $2^{j \cdot (j-1)} \cdot (j!)^2$ and, hence, by $2^{2(k-1) \cdot (2k-3)} \cdot ((2k-2)!)^2$. Thus, Eq (3.18) goes as

$$(3.19) \quad \delta_2 \leq \frac{c_k}{\mu \cdot M^2} \sum_{j=k}^{2(k-1)} \sum_{(s_1, T_1) \in A} \sum_{(s_3, T_3) \in U(s_1, T_1, j)} \frac{1}{\psi(s_3, T_3, D)} E \left[\mathbf{1}_{s_3}(T_3, \tilde{\mathcal{G}}) \right],$$

where c_k is a general constant. Let us utilize the same argument as in Lemma 3.3 to see that,

$$(3.20) \quad \begin{aligned} \delta_2 &\leq \frac{c_k}{\mu \cdot M^2} \sum_{j=k}^{2(k-1)} \sum_{(s_1, T_1) \in A} \sum_{(s_3, T_3) \in U(s_1, T_1, j)} E \left[\frac{1}{\psi(s_3, T_3, \tilde{\mathcal{G}})} \mathbf{1}_{s_3}(T_3, \tilde{\mathcal{G}}) \right] \\ &+ \frac{c_k}{\mu \cdot M^2} M^a n^{1-a}. \end{aligned}$$

For the first term in the above equation, we use an idea similar to Lemma 3.1. So we obtain the following uniform bound for $k \leq j \leq 2(k-1)$,

$$\begin{aligned} F(s_1, T_1, G) : &= \sum_{(s_3, T_3) \in U(s_1, T_1, j)} \frac{1}{\psi(s_3, T_3, G)} \mathbf{1}_{s_3}(T_3, G), \\ &\leq \frac{1}{\psi(s_1, T_1, G)} \mathbf{1}_{s_1}(T_1, G). \end{aligned}$$

In addition, the second term in eq (3.20) is bounded by $\frac{c_k}{\mu \cdot M}$, because $n < M$ and $a < 1$. Therefore,

$$\begin{aligned} \delta_2 &\leq \frac{c_k}{\mu \cdot M^2} \sum_{j=k}^{2(k-1)} \sum_{(s_1, T_1) \in A} E \left[\frac{1}{\psi(s_1, T_1, \tilde{\mathcal{G}})} \mathbf{1}_{s_1}(T_1, \tilde{\mathcal{G}}) \right] \\ &\leq \frac{c_k}{\mu \cdot M^2} \sum_{(s_1, T_1) \in A} E \left[\frac{1}{\psi(s_1, T_1, \tilde{\mathcal{G}})} \mathbf{1}_{s_1}(T_1, \tilde{\mathcal{G}}) \right], \end{aligned}$$

where we put $2(k-1) - k = k - 2$ into the general constant. Again, we change $\psi(s_1, T_1, \tilde{\mathcal{G}})$ back to $\psi(s_1, T_1, D)$, where by Lemma 3.3, it costs us another $c_k M^a n^{1-a} \leq c_k M$. Thus, it follows by Eq. (3.17) that

$$\begin{aligned} \delta_2 &\leq \frac{c_k}{\mu \cdot M^2} \left[\sum_{(s_1, T_1) \in A} \frac{1}{\psi(s_1, T_1, D)} E \left[\mathbf{1}_{s_1}(T_1, \tilde{\mathcal{G}}) \right] + M \right] \\ &\leq \frac{c_k}{M}. \end{aligned}$$

This upper bound and Eq. (3.16), along with Theorem A.1, complete the proof of this preposition. \square

Now, we can prove Theorem 2.11.

Proof of Theorem 2.11. The first step is to split the trees into two sets, A^+ and A^- , and show that it suffices to prove the statement for one of the sets like A^- . We let

$$A^+ := \left\{ (s, T) \in \mathbb{S}^k \times \mathbb{T}^k \mid \mathbf{p}_a(s, T) - \tilde{\mathbf{p}}(s, T) \geq 0 \right\},$$

and

$$\mu^+ := \sum_{(s, T) \in A^+} \frac{1}{\psi(s, T, D)} \tilde{\mathbf{p}}(s, T) = \sum_{(s, T) \in A^+} \frac{1}{\psi(s, T, D)} E \left[\mathbf{1}_s(s, T) \right],$$

and

$$A^- := \{(s, T) \in \mathbb{S}^k \times \mathbb{T}^k \mid \mathbf{p}_{\mathbf{a}}(s, T) - \tilde{\mathbf{p}}(s, T) < 0\},$$

and we define μ^- similarly. Theorem 2.16 states that

$$\mu^+ + \mu^- = \sum_{(s, T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} \tilde{\mathbf{p}}(s, T) = k^{k-2} + O\left(\left(\frac{M}{n}\right)^{-\frac{1}{2}}\right).$$

We let

$$I^+ := \sum_{(s, T) \in A^+} \frac{1}{\psi(s, T, D)} (\mathbf{p}_{\mathbf{a}}(s, T) - \tilde{\mathbf{p}}(s, T)),$$

and I^- correspondingly. Combining part 3 of Theorem 2.16 with the above equation, we get

$$I^+ = I^- + O\left(\left(\frac{M}{n}\right)^{1-a}\right).$$

Thus,

$$\begin{aligned} L_{\mathbf{a}}(a, k, D) &:= \sum_{(s, T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} |\mathbf{p}_{\mathbf{a}}(s, T) - \tilde{\mathbf{p}}(s, T)| \\ (3.21) \quad &= 2I^- + O\left(\left(\frac{M}{n}\right)^{1-a}\right). \end{aligned}$$

We let $a_1 := a - \frac{1}{2}$ and, by the hypothesis of the theorem, we get $0 < a_1 < \frac{\nu}{4}$. We also recall from Assumption 2.7 that $\left(\frac{M}{n}\right)^{-\frac{1}{2}} < n^{-\nu}$. Taking into account that $M < n^2$, it follows that $\left(\frac{M}{n}\right)^{1-a} \leq n^{-\nu+a_1}$. Now, we may assume that $\mu^- \geq n^{-\nu+3a_1}$, otherwise

$$I^- \leq \sum_{(s, T) \in A^-} \frac{1}{\psi(s, T, D)} \tilde{\mathbf{p}}(s, T) = \mu^- < n^{-\nu+3a_1} = O\left(n^{a_1-\frac{\nu}{2}}\right),$$

and nothing remains to be proven. All of the above considerations are meant to result in a bound on ϵ , which is defined as

$$(3.22) \quad \epsilon^2 := \frac{n^{-\nu+2a_1}}{\mu^-} \leq n^{-a_1} \ll 1.$$

Next, we change the underlying measure of $\mathcal{G}_{\mathbf{a}}(a, D)$ to $\tilde{\mathcal{G}}(D)$ with the use of Propositions 3.6 and 3.4. Then, we apply Proposition 3.9. Hence, for $0 < \epsilon \ll 1$,

$$\begin{aligned} (3.23) \quad &L^- \\ &:= P(|I^-| > \mu^- \epsilon) \\ &= P\left(\left|\sum_{(s, T) \in A^+} \frac{1}{\psi(s, T, D)} (\tilde{\mathbf{p}}(s, T) - \mathbf{1}_s(T, \mathcal{G}_{\mathbf{a}}(a, D)))\right| > \mu^- \epsilon\right) \\ &\leq 2e^{10\mathcal{C}_1(D)+2\mathcal{C}_2(D)} P\left(\left|\sum_{(s, T) \in A^+} \frac{1}{\psi(s, T, D)} (\tilde{\mathbf{p}}(s, T) - \mathbf{1}_s(T, \tilde{\mathcal{G}}(D)))\right| > \mu^- \epsilon\right) \\ &\leq 2 \exp(10\mathcal{C}_1(D) + 2\mathcal{C}_2(D) - (c\mu^- M) \epsilon^2). \end{aligned}$$

The ϵ in (3.22) minimizes the right hand side of the above equation. So, $\epsilon^2 = \frac{n^{-\nu+2a_1}}{\mu^-}$, and we bound $\mathcal{C}_1(D)$ and $\mathcal{C}_2(D)$ according to Lemma 3.8. Thus,

$$\begin{aligned} L^- &\leq \exp(c \cdot n^{a_1-\nu} \cdot M \cdot \log(n) - c \cdot n^{2a_1-\nu} \cdot M) \\ &= \exp(-c \cdot n^{2a_1-\nu} \cdot M) \left(1 - O\left(\frac{\log(n)}{n^{a_1}}\right)\right) \end{aligned}$$

that goes to zero faster than any polynomial in n , since a_1 is positive. So there exists an N such that, for all $n > N$, we get $L^- \mu^- (1 - \epsilon) \leq L^- n^2 < 1$.

If we define

$$F := \left\{ G \in \mathbb{G}_a^D : \sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} \mathbf{1}_s(T,G) < \left[\sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} \tilde{\mathbf{p}}(s,T) \right] (1 - \epsilon) \right\}$$

then,

$$\begin{aligned} &\sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} \mathbf{p}_a(s,T) \\ &= E \left[\sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} \mathbf{1}_s(T, \mathcal{G}_a) \right] \\ &\geq 0 \cdot P(F) + \left(\sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} \tilde{\mathbf{p}}(s,T) (1 - \epsilon) \right) \cdot P_a(F^c) \\ &\geq 0 + (1 - L^-) \mu^- (1 - \epsilon) \\ &= \mu^- - \mu^- \epsilon - L^- \mu^- (1 - \epsilon). \end{aligned}$$

It follows from part 3 of Theorem 2.16 that $\mu^- \leq k^{k-2} + 1 \leq c_k$ for large n . Therefore, using (3.22) we obtain

$$\begin{aligned} \sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} (\tilde{\mathbf{p}}(s,T) - \mathbf{p}_a(s,T)) &= \mu^- - \sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} \mathbf{p}_a(s,T) \\ &\leq \mu^- \epsilon - 1 \\ &< (\mu^- \cdot n^{-\nu+2a_1})^{\frac{1}{2}} \\ &\leq c_k \cdot n^{-\frac{\nu}{2}+a_1}, \end{aligned}$$

and

$$L_a(a, k, D) = I^- + I^- \leq c_k \cdot (n^{-\frac{\nu}{2}+a_1} + n^{-\nu+a_1}) \leq c_k \cdot n^{-\frac{\nu}{2}+a-\frac{1}{2}}.$$

That does it. \square

We assumed that the vector D is a strict graphic sequence of type (ε, ν) (Definition 2.8) and, as we saw, that provides some bounds on numbers r_i s. Although, those bounds were crucial for our main proof, we believe that they hold in wider generality.

Conjecture 3.13. *Suppose that D is a strict graphic sequence as in Definition 2.3. Also, for $1 \leq i \leq n$, let the variables r_i s be defined as in Eq. (3.3), then*

$$|\log(r_i)| < c \cdot \log(n),$$

where $c > 0$.

Remark 3.14. Lemma 3.8 gives the rough bound of $O\left(n^{a-\frac{1}{2}-\nu} \cdot M \cdot \log(n)\right)$ on $\mathcal{C}_2(D)$ for a strict graphic sequence D of type (ε, ν) . However, if we believe that r_i s have polynomial bounds in n , as in Conjecture 3.13, then we get a better bound. Indeed, we obtain

$$\begin{aligned} \mathcal{C}_2(D) &= \sum_{i=1}^n d_i^a |\log(r_i)| \\ &= O\left(\log(n) \left(\sum_{i=1}^n d_i^a\right)\right) \\ (3.24) \quad &= O\left(\log(n) \cdot n^{a-\frac{1}{2}} \cdot M \cdot \left(\frac{n}{M}\right)^{1/2}\right), \end{aligned}$$

where we bound d_i by n , for $i \in [n]$ and use Cauchy-Schwarz for $\sum_{i=1}^n d_i^{\frac{1}{2}}$. Note that $\left(\frac{n}{M}\right)^{1/2}$ can be much smaller than $n^{-\nu}$, which can lead to a better bound in Theorem 2.11. Actually, a careful investigation of our previous proof shows that if we pick a such that $n^{4(a-\frac{1}{2})} \ll \left(\frac{n}{M}\right)^{1/2}$, and if Conjecture 3.13 holds, then

$$L_{\mathbf{a}}(a, k, D) \leq c_k \left(\sqrt{\frac{n}{M}}\right)^{\frac{1}{2}} n^{(a-\frac{1}{2})} = c_k \left(\frac{n}{M}\right)^{\frac{1}{4}} n^{(a-\frac{1}{2})}.$$

Remark 3.15. Conjecture 3.13 holds if $d_n^2 < \frac{1}{2}M$. Indeed, let us write

$$\sum_{i=1}^{n-d_n} d_i = M - \sum_{i=n-d_n+1}^n d_i \geq M - d_n^2 > \frac{1}{2}M.$$

Using part 5 of previous lemma, we have $r_n \leq n^2$. Also, from part 1 of that lemma, $r_1 \geq n^{-3}$. Therefore, $|\log(r_i)| \leq 3\log(n)$.

3.2. A lower bound. Our goal is to prove Proposition 3.4, that is

$$P\left(\tilde{\mathcal{G}}(D) \in \mathbb{G}_a^D\right) > \exp(-10 \cdot \mathcal{C}_1(D)),$$

where

$$\mathcal{C}_1(D) := |\log(\delta)| \cdot n \cdot \log^{\frac{10}{a_1}}(n),$$

and δ is such that $\delta \leq \tilde{p}_{ij} \leq 1 - \delta$, and $a_1 := a - \frac{1}{2}$.

Throughout this section, we use the following notations frequently.

Definition 3.16. For the sets $A, B \subseteq [n]$, we define $Ed(A, B)$ as the set of all edges with one end in A and the other end in B , or

$$Ed(A, B) := \{\langle i, j \rangle \mid i \in A, j \in B\}.$$

In addition, we use the short version $Ed(A)$ for $Ed(A, A)$.

Correspondingly, $Ed([n])$ is the set of all possible edges on $[n]$.

Definition 3.17. Let $E \subseteq \text{Ed}([n])$ be a collection of edges. We define $\tilde{\mathcal{G}}(E, D)$ as the restriction of $\tilde{\mathcal{G}}(D)$ to the edge set E , i.e. $\tilde{\mathcal{G}}(E, D)$ is a random graph with independent Bernoulli edges with probability \tilde{p}_{ij} , where $\langle i, j \rangle \in E$. We also show random graphs $\tilde{\mathcal{G}}(\text{Ed}(A), D)$, $\tilde{\mathcal{G}}(\text{Ed}(A, B), D)$, and $\tilde{\mathcal{G}}(\text{Ed}(B), D)$, with $\tilde{\mathcal{G}}_A$, $\tilde{\mathcal{G}}_{A,B}$, and $\tilde{\mathcal{G}}_B$, respectively.

It is easy to check that the following theorem implies Proposition 3.4, and we spend the rest of this section proving that.

Theorem 3.18. *We use the notation in Def. (3.16) and (3.17), and $0 < a < 1$, and $a_1 = a - \frac{1}{2}$ are constants independent of n . In addition, we assume that the \tilde{p}_{ij} s in 3.17, for $i, j \in [n]$, satisfy $\delta \leq \tilde{p}_{ij} \leq 1 - \delta$. Then, there is a partition of $[n]$ into two sets A and B , and there exists a deterministic tree T with edges in $\text{Ed}(A, B)$ such that, for all large enough n ,*

a)

$$P\left(\tilde{\mathcal{G}}_A = \emptyset_A\right) \geq e^{-4 \cdot |\log(\delta)| \cdot n \cdot \log^{\frac{10}{a_1}}(n)},$$

b) and

$$P\left(\tilde{\mathcal{G}}_{A,B} = T\right) \geq e^{-5 \cdot |\log(\delta)| \cdot n \cdot \log^{\frac{10}{a_1}}(n)},$$

c) and

$$P\left(\emptyset_A \cup T \cup \tilde{\mathcal{G}}_B \in \mathbb{G}_a^D\right) \geq \frac{1}{2}.$$

Here $G_1 \cup G_2$ is the union of edges of the two graphs G_1 and G_2 , and the graph \emptyset_A is a graph on vertices of A with no edges.

We actually prove the above theorem for $A = I_{\frac{10}{a_1}}$ and $B = [n] \setminus I_{\frac{10}{a_1}}$, where I_α , for any $\alpha > 0$, is defined as

$$(3.25) \quad I_\alpha := \{i \mid d_i \leq \log^\alpha(n)\}.$$

Our proof has two steps. First, we build the deterministic graph T , and then we show that it can be extended to a graph with an almost given degree sequence with high probability.

Lemma 3.19. *Let us define*

$$D(i, A) := \sum_{j \in A \setminus \{i\}} \tilde{p}_{ij},$$

where $i \in [n]$ and A is a subset of $[n] = \{1, \dots, n\}$. Then, for every $A \subseteq [n]$, there is a bipartite graph T with vertex sets A and $B = [n] \setminus A$, and edges inside $\text{Ed}(A, B)$ such that

a)

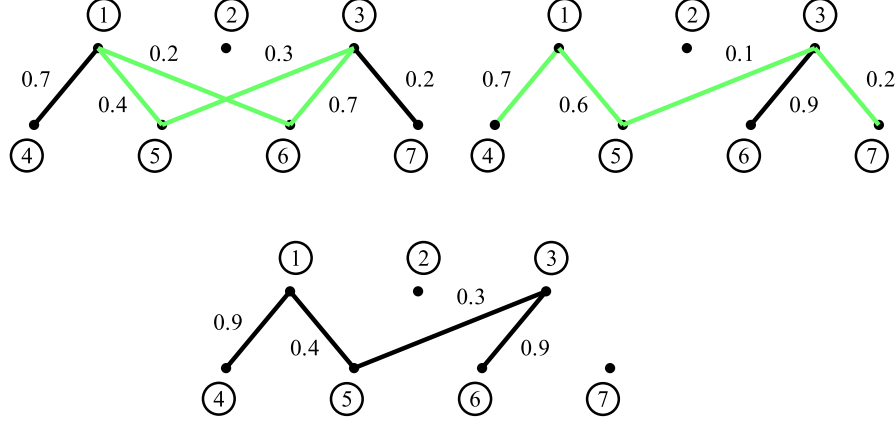
$$(3.26) \quad b_i \in \{\lfloor D(i, B) \rfloor, \lfloor D(i, B) \rfloor + 1\}, \text{ where } i \in A,$$

b) and,

$$(3.27) \quad b_i \in \{\lfloor D(i, B) \rfloor, \lfloor D(i, B) \rfloor + 1\}, \text{ where } i \in B,$$

where b_i is the degree of vertex i in the graph T and $\lfloor x \rfloor$ is the biggest integer less than x , for $x \in \mathbb{R}$.

FIGURE 3.1. The process of changing a weighted bipartite graph to a bipartite graph with weights in $\{0, 1\}$.



The figure on the upper left hand side shows the live edges of a bipartite graph on 7 vertices. We pick a loop that is shown in green, and change the weights on it.

We add 0.2 to the weights of edges $\langle 1, 5 \rangle$ and $\langle 3, 6 \rangle$, and subtract 0.2 from $w(\langle 5, 3 \rangle)$ and $w(\langle 6, 1 \rangle)$. Now, the edge $\langle 1, 6 \rangle$ dies and we get the figure on the upper right hand side. Since we have no other loops, we choose the biggest path that is shown in green. If we add and subtract 0.1 to and from the weights of edges of this path alternatively, we get the third graph, where the edge $\langle 3, 7 \rangle$ becomes dead.

Proof. We already have a weighted bipartite graph T with edge weights $0 \leq \tilde{p}_{ij} \leq 1$ that almost satisfies the conditions (3.26) and (3.27). Our goal is to change the weights continuously to get weights of size 0 or 1 without changing the degrees of our graph as much as possible. Let us denote the weight of an edge e by $w(e)$. In addition, we say an edge e is dead if $w(e) \in \{0, 1\}$, and is live if $0 < w(e) < 1$.

Let us start a two stage process, where an example of that is shown at Fig. 3.1. We pick a closed loop of live edges, namely e_1, \dots, e_{2k} . Note that such a loop has an even number of edges since the graph T is bipartite. Now, we add and subtract a constant number c alternatively to the weights of e_i , for $1 \leq i \leq k$, to get new weights $w(e_1) + c, w(e_2) - c, \dots$, and $w(e_k) + (-1)^{k-1}c$. Let c grow gradually from 0 until the first edge dies. Since the loop has an even length, the degrees of the graph have not changed. We keep on doing this until all loops disappear.

Next, we pick a path from the longest live paths in T . Suppose that the path runs through vertices v_1, \dots, v_k . We observe that if v_1 is attached to two live edges, we can either make our path longer or we get a loop amongst vertices v_i s. Since none of them are possible, v_1 is only attached to one live edge, and the same is true for v_k , the last vertex of the path. Again, we change the weights alternatively by $+c$ and $-c$, to get the new weights as $w(\langle v_1, v_2 \rangle) + c, w(\langle v_2, v_3 \rangle) - c, \dots$, and $w(\langle v_{k-1}, v_k \rangle) + (-1)^{k-1}c$. Then, we let c grow gradually until an edge dies. We repeat this process until no live edges are left.

Since the dead edges have a weight of 0 or 1, we have reached a bipartite graph. It remains to be shown that our new graph satisfies conditions (3.26) and (3.27). As we saw before, the degrees of vertices of T do not change during the first stage

of the process. Similarly, the degrees do not change in the second part, except for the end vertices of the paths. In addition, as we discussed earlier, if a vertex v becomes an end point for a path in our procedure, that vertex is only attached to one live edge. Therefore, its degree $d(v)$ has changed by at most the amount of changes on the weight of that edge. Hence, $d(v)$ has become $\lfloor d(v) \rfloor$ or $\lfloor d(v) \rfloor + 1$, which completes the proof. \square

Lemma 3.20. *If $\delta \leq \tilde{p}_{ij} \leq 1 - \delta$, for some $\delta > 0$, and $1 \leq i, j \leq n$, then,*

$$P\left(\tilde{\mathcal{G}}_A = \emptyset_A\right) \geq e^{-5 \cdot |\log(\delta)| \cdot n \log^\alpha(n)},$$

where $A = I_\alpha$ and $B = [n] \setminus I_\alpha$, for some positive α (in particular $\alpha = \frac{10}{a_1}$), and I_α is as in Eq. (3.25). In addition, for the bipartite graph $T \subseteq Ed(A, B)$ in the previous lemma,

$$P\left(\tilde{\mathcal{G}}_{A,B} = T\right) \geq e^{-5 \cdot |\log(\delta)| \cdot n \log^\alpha(n)}.$$

Proof. We have $d_i \leq \log^\alpha(n)$, for $i \in A$, and $|A| \leq n$. Therefore,

$$(3.28) \quad S_1 := \sum_{i \in A} D(i, A) = \sum_{i \in A} \sum_{j \in A, j \neq i} \tilde{p}_{ij} \leq \sum_{i \in A} d_i \leq n \log^\alpha(n).$$

Let F_1 be the set of edges $\langle i, j \rangle$ in $Ed(A)$ such that $\tilde{p}_{ij} > \frac{1}{2}$. We see that $|F_1| \leq 2n \log^\alpha(n)$, otherwise, S_1 would exceed the right hand side of Eq. (3.28). Hence,

$$\begin{aligned} P(\tilde{\mathcal{G}}_A = \emptyset_A) &= \prod_{\langle i, j \rangle \in Ed(A)} (1 - \tilde{p}_{ij}), \\ &= \prod_{\langle i, j \rangle \in F_1} (1 - \tilde{p}_{ij}) \cdot \prod_{\langle i, j \rangle \in Ed(A) \setminus F_1} (1 - \tilde{p}_{ij}) \\ &\geq \delta^{|F_1|} \prod_{\langle i, j \rangle \in Ed(A) \setminus F_1} e^{-2\tilde{p}_{ij}} \\ &\geq \delta^{|F_1|} e^{-2S_1} \\ &\geq e^{-(2 \cdot |\log(\delta)| + 2) \cdot n \log^\alpha(n)}, \end{aligned}$$

where we have used the inequality $1 - x \geq e^{-2x}$, for $x \leq \frac{1}{2}$, and Eq. (3.28). This concludes the first part, since $\delta \leq \frac{1}{2}$ and $3 \log(2) \geq 3 |\log(\delta)| > 2$.

For the second part, we define $F_2 \subseteq Ed(A, B)$ similar to the set F_1 , and

$$(3.29) \quad S_2 := \sum_{i \in A} D(i, B) = \sum_{i \in A} \sum_{j \in B} \tilde{p}_{ij} \leq \sum_{i \in A} d_i \leq n \log^\alpha(n).$$

Let E be the set of edges of the graph T . Then,

$$\begin{aligned} P(\tilde{\mathcal{G}}_{A,B} = T) &= \prod_{\langle i, j \rangle \in E} \tilde{p}_{ij} \cdot \prod_{\langle i, j \rangle \in Ed(A,B) \setminus E} (1 - \tilde{p}_{ij}), \\ &\geq \delta^{|E|} \cdot \prod_{\langle i, j \rangle \in F_2} (1 - \tilde{p}_{ij}) \prod_{\langle i, j \rangle \in Ed(A,B) \setminus (E \cup F_2)} (1 - \tilde{p}_{ij}) \\ &\geq \delta^{|E| + |F_2|} \cdot \prod_{\langle i, j \rangle \in Ed(A,B) \setminus F_2} e^{-\tilde{p}_{ij}} \\ &\geq \delta^{|E| + |F_2|} e^{-2S_2}. \end{aligned}$$

Let us observe from the previous lemma that $|E|$, the number of edges of T , is less than or equal to $\frac{1}{2}S_2 + n$. Using a bound on F_2 similar to that of F_1 , and by (3.29) we get our result. \square

In order to get the third part of Theorem (3.18), we use the union bound on the random graph $\tilde{\mathcal{G}}_B$, where $B = [n] \setminus I_{\frac{10}{\alpha_1}}$. Hence, the following proposition is handy.

Proposition 3.21. *For every $\alpha, \beta \in \mathbb{R}^+$, and large enough n (depending on α and β), we have,*

$$D(j, I_\alpha) < \frac{1}{4},$$

where $j \in I_\beta$, and I_α is as in Eq. (3.25).

Proof. We prove this through a series of lemmas. Let us recall from the proof of Proposition 2.1 that, for the degree sequence $d_1 \leq d_2 \leq \dots \leq d_n$, there exist positive real numbers $0 < r_1 \leq \dots \leq r_n$ such that

$$\tilde{p}_{ij} = \frac{r_i r_j}{1 + r_i r_j},$$

and,

$$d_i = \sum_{j \in [n] \setminus \{i\}} \tilde{p}_{ij},$$

where $[n] = \{1, \dots, n\}$. Also recall that $M = \sum_{i \leq n} d_i > n^{1+\epsilon}$, for large enough n . Moreover, let us assume that $I_\alpha = \{1, \dots, k(\alpha)\}$, where $k(\alpha)$ is the number of elements of I_α , and

$$(3.30) \quad L_\alpha := \{i : i \leq \ell(\alpha)\} \quad \text{with} \quad \ell(\alpha) = n - 2d_{k(\alpha)}.$$

We note that $k(\alpha) \in I_\alpha$, and hence, $d_{k(\alpha)} \leq \log^\alpha(n)$. In addition, for large enough n ,

$$\begin{aligned} \sum_{i=1}^{k(\alpha)} d_i + \sum_{i=\ell}^n d_i &\leq k(\alpha) \cdot d_{k(\alpha)} + (2d_{k(\alpha)} + 1) \cdot n \\ &\leq n \cdot \log^\alpha(n) + (2 \log^\alpha(n) + 1) \cdot n \\ &\leq n^{1+\epsilon} < M. \end{aligned}$$

Therefore, $k(\alpha) < \ell(\alpha)$.

Now that the dependency of I_α , L_α , $k(\alpha)$, and $\ell(\alpha)$ on α is understood, we drop the α for the convenience of our notation, and use I , L , k and ℓ throughout the rest of our proof.

Lemma 3.22. *For $j \in I$,*

$$D(j, I) \leq 2d_j \frac{\sum_{i \in I} r_i}{\sum_{i \in L} r_i}.$$

Proof. We see that $\tilde{p}_{ij} = \frac{r_i r_j}{1 + r_i r_j}$ is increasing both in i and j since r_i s are increasing in i , for $1 \leq i, j \leq n$. Thus, it follows by Eq. (3.30) that

$$d_k \geq \sum_{j \geq \ell+1} \tilde{p}_{kj} \geq 2d_k \tilde{p}_{k\ell}.$$

This implies $\tilde{p}_{kl} \leq \frac{1}{2}$ and therefore, for $i \leq k$ and $j \leq \ell$, we get $\tilde{p}_{ij} \leq \frac{1}{2}$ and $r_i r_j \leq 1$. We can now estimate for $j \leq \ell$,

$$D(j, I) = \sum_{i \leq k} \tilde{p}_{ij} = \sum_{i \leq k} \frac{r_i r_j}{1 + r_i r_j} \leq r_j \sum_{i \leq k} r_i.$$

Furthermore,

$$r_j = \frac{d_j r_j}{d_j} = \frac{d_j r_j}{\sum_{1 \leq i \leq n} \frac{r_i r_j}{1 + r_i r_j}} \leq \frac{d_j r_j}{\sum_{1 \leq i \leq \ell} \frac{r_i r_j}{1 + r_i r_j}} \leq \frac{2d_j r_j}{\sum_{1 \leq i \leq \ell} r_i r_j} \leq \frac{2d_j}{\sum_{1 \leq i \leq \ell} r_i}.$$

That does it. \square

Lemma 3.23. *We have,*

$$\sum_{i \leq k} r_i \leq \sqrt{(4n+2)d_k}.$$

Proof. Note that as in the previous lemma $r_i r_j \leq 1$, for $i \leq k$ and $j \leq \ell$. Then, we observe that $k < \ell$, and

$$\begin{aligned} \left[\sum_{i \leq k} r_i \right]^2 &= 2 \sum_{i < j \leq k} r_i r_j + \sum_{i \leq k} r_i^2 \\ &\leq 4 \sum_{i < j \leq k} \frac{r_i r_j}{1 + r_i r_j} + \sum_{i \leq k} r_i r_k \\ &\leq 4 \sum_{i < j \leq k} \frac{r_i r_j}{1 + r_i r_j} + 2 \sum_{i \leq k} \frac{r_i r_k}{1 + r_i r_k} \\ &\leq 4 \sum_{i \leq k} d_i + 2d_k \\ &\leq (4n+2)d_k. \end{aligned}$$

\square

Lemma 3.24. *For large enough n , we get*

$$\sum_{i \leq \ell} r_i \geq \sqrt{\frac{M}{2}}.$$

Proof. The proof goes through the following lines

$$\left[\sum_{i \leq \ell} r_i \right]^2 \geq \sum_{1 \leq i, j \leq \ell} \frac{r_i r_j}{1 + r_i r_j} \geq \sum_{1 \leq i, j \leq \ell} \tilde{p}_{ij} \geq M - 2 \sum_{i > \ell} d_i \geq M - 4nd_k.$$

Since $d_k \leq \log^\alpha(n)$, and $M > n^{1+\epsilon}$, we get our result for large enough n . \square

Let us go back to the proof of our proposition. Since $k \in I_\alpha$ and $j \in I_\beta$, we have $d_j \leq \log^\beta(n)$ and $d_k \leq \log^\alpha(n)$. In addition, $M > n^{1+\epsilon}$, for large values of n . Combining the above three lemmas, we obtain

$$D(j, I) \leq 2d_j \sqrt{\frac{2(4n+2)d_k}{M}} \leq 8 \frac{\log^{\beta+\frac{\alpha}{2}}(n)}{n^{\frac{\epsilon}{2}}} < \frac{1}{4},$$

for large enough n . \square

We continue with the last part of Theorem 3.18. Recall that $\tilde{\mathcal{G}}_B$ is the random graph with bernoulli random edges with parameter \tilde{p}_{ij} , for $i, j \in B$, where $B = [n] \setminus I_{\frac{10}{a_1}}$ and $0 < a_1 = a - \frac{1}{2} < \frac{1}{2}$.

Lemma 3.25. *Let $\mathbf{1}_{\langle i, j \rangle \in \mathcal{E}(G)}$ for $i, j \in B$, be the indicator of the edge $\langle i, j \rangle$ in graph G . Then, we define the events,*

$$(3.31) \quad E_j := \left\{ G \text{ such that } \left| \sum_{i \in B} \mathbf{1}_{\langle i, j \rangle \in \mathcal{E}(G)} - D(j, B) \right| \leq D(j, B)^a \right\},$$

and

$$(3.32) \quad F_j := \left\{ G \text{ such that } \sum_{i \in B} \mathbf{1}_{\langle i, j \rangle \in \mathcal{E}(G)} \leq (2 \log^2(n) + 1) D(j, B) \right\},$$

for $j \in B$. In addition, for $0 < a_1 = a - \frac{1}{2} < \frac{1}{2}$, we define

$$(3.33) \quad J := \left\{ j \in [n] \mid D(j, B) \geq \log^{\frac{1}{a_1}}(n) \right\}.$$

Then, for large enough n ,

$$P\left(\tilde{\mathcal{G}}_B \in \cap_{j \in J} E_j \cap_{j \notin J} F_j\right) \geq 1 - \frac{1}{n} > \frac{1}{2}.$$

Proof. Let X_1, \dots, X_n be a vector of independent Bernoulli random variables with total mean $\mu := \sum_{1 \leq i \leq n} E[X_i]$. The Chernoff's bound [15] states that, for $\delta > 0$,

$$P\left(\sum_{1 \leq i \leq n} X_i \geq (1 + \delta)\mu\right) \leq \exp\left(\frac{-\mu \cdot \delta^2}{2 + \delta}\right),$$

and

$$P\left(\sum_{1 \leq i \leq n} X_i \leq (1 - \delta)\mu\right) \leq \exp\left(\frac{-\mu \cdot \delta^2}{2 + \delta}\right).$$

We apply Chernoff's bound with parameters

$$\mu_j := \sum_{i \in B \setminus \{j\}} E[\mathbf{1}_{\langle i, j \rangle}] = \sum_{i \in B \setminus \{j\}} \tilde{p}_{ij} = D(j, B),$$

and $\delta = D(j, B)^{-\frac{1}{2} + a_1} \leq 1$, where $1 \leq j \leq n$. Hence,

$$(3.34) \quad P(E_j^c) \leq 2 \exp\left(-\frac{1}{3} D(j, B)^{2a_1}\right).$$

Now, for those j in J and by (3.33), we observe that Eq. (3.34) turns into

$$(3.35) \quad P(E_j^c) \leq 2e^{-\frac{1}{3} \log^2(n)}.$$

Similarly, for $1 \leq j \leq n$,

$$(3.36) \quad P(F_j^c) \leq \exp(-\log^2(n) D(j, B)),$$

where $\delta = 2 \log^2(n) D(j, B)$. In order to complete the bound in Eq (3.36), we show $D(j, B) \geq \frac{1}{4}$, for $j \in B$. If $A = I_{\frac{10}{a_1}} = [n] \setminus B$ is empty then $D(j, B) = d_j \geq 1$.

Otherwise, if $A \neq \emptyset$, then $1 \in A$. We observe that $\frac{r_i r_j}{1+r_i r_j}$ is increasing both in i and j , so

$$\begin{aligned}
 (3.37) \quad D(j, B) &= \sum_{i \in B \setminus \{j\}} \tilde{p}_{ij} = \sum_{i \in B \setminus \{j\}} \frac{r_i r_j}{1+r_i r_j} \\
 &\geq \sum_{i \in B} \left(\frac{r_i r_1}{1+r_i r_1} \right) - \frac{r_1 r_j}{1+r_1 r_j} \\
 &= D(1, B) - \tilde{p}_{1j} \\
 &= d_1 - D(1, A) - \tilde{p}_{1j}.
 \end{aligned}$$

In addition, we know that $d_1 \geq 1$, and by Proposition 3.21, we get $D(1, A) \leq \frac{1}{4}$, for large enough n . Thus, if \tilde{p}_{1j} is smaller than $\frac{1}{4}$, then Eq. (3.37) gives $D(j, B) \geq \frac{1}{2}$. On the other hand, we get $\frac{1}{4} \leq \tilde{p}_{1j} \leq D(j, B)$.

Therefore,

$$P(F_j^c) \leq e^{-\frac{1}{4} \log^2(n)}.$$

Combining the previous inequality with Eq. (3.35), we have

$$P\left(\tilde{\mathcal{G}}_B \in \cap_{j \in J} E_j \cap_{j \notin J} F_j\right) \geq 1 - \sum_{i \in J} e^{-\frac{1}{3} \log^2(n)} - \sum_{i \notin J} e^{-\frac{1}{4} \log^2(n)} \geq 1 - \frac{1}{n},$$

for large enough n . That concludes the proof. \square

Lemma 3.26. *Let E_j , F_j , and J be as they are in Lemma 3.25. In addition, T is the resulting bipartite graph in Lemma 3.19. Then, for every graph $G \in \cap_{j \in J} E_j \cap_{j \notin J} F_j$, we have*

$$\emptyset_A \cup T \cup G \in \mathbb{G}_a^D,$$

where $A = I_{\frac{10}{a-1}}$, and the graph \emptyset_A is a graph on vertices of A with no edges.

Proof. Let us denote $d_i(G)$ by the degree of the i^{th} vertex of the graph G . As usual, B is $[n] \setminus A$. We need to show that

$$s_i := |d_i - (d_i(\emptyset_A) + d_i(T) + d_i(G))| \leq 2d_i^a,$$

for every $i \in [n]$. For $i \in A$, it follows by Lemma 3.19 and Proposition 3.21 that, for large enough n ,

$$s_i = |d_i - d_i(T)| \leq 1 + |d_i - D(i, B)| = D(i, A) + 1 < 2.$$

Again, we let J be the same as in (3.33). In addition, for $i \in J \cap B$, we use the definition of E_j (Eq. (3.31)), to get

$$\begin{aligned}
 s_i &= |d_i - d_i(T) - d_i(G)| \\
 &\leq 1 + |d_i - D(i, A) - d_i(G)| \\
 &= 1 + |D(i, B) - d_i(G)| \\
 &\leq 1 + D(i, B)^a \\
 &\leq 2d_i^a.
 \end{aligned}$$

Last, for $i \in B \setminus J$, and using the property of the set F_i (Eq. (3.32)), we obtain,

$$\begin{aligned} s_i &= |d_i - d_i(T) - d_i(G)| \\ &\leq 1 + |D(i, B) - d_i(G)| \\ &\leq 1 + D(i, B) + (2 \log^2(n) + 1) D(i, B) \\ &\leq 1 + 2 (\log^2(n) + 1) D(i, B). \end{aligned}$$

Note that $i \notin J$, so $D(i, B) \leq \log^{\frac{1}{a_1}}(n)$, where $a_1 = a - \frac{1}{2}$. In addition, $i \in B = [n] \setminus I_{\frac{10}{a_1}}$, where $I_{\frac{10}{a_1}}$ is defined by Eq. (3.25). Therefore, $d_i \geq \log^{\frac{10}{a_1}}(n)$, and

$$s_i \leq 8 \log^{2+\frac{1}{a_1}}(n) \leq 2d_i^a,$$

where we used $a < 1$. This completes the proof. \square

Proof of Theorem 3.18. As we saw before, the partition is $A = I_{\frac{10}{a_1}}$, and $B = [n] \setminus A$, and $I_{\frac{10}{a_1}}$ is as in Eq. (3.25). Lemma 3.19 shows us that there exists a bipartite tree with edges in $Ed(A, B)$. From Lemma 3.20, we get parts a and b of the theorem.

Finally, putting Lemmas 3.25 and 3.26 together, we get the required lower bound for part c of the theorem. \square

3.3. Graphs with a given degree sequence. Throughout this section, we let C be a general constant. In addition, most of the proofs are analogous to the proof of Theorem 2.11 with some changes. So we provide an outline for each solution as well as the essential steps.

Recall that

$$\mathbf{p}_{\mathbf{g}}(s, T) = E[\mathbf{1}_s(T, \mathcal{G}_{\mathbf{g}})],$$

where $\mathcal{G}_{\mathbf{g}}$ is the random graph chosen uniformly in \mathbb{G}^D .

Corollary 3.27. *If we use the notation in Conjecture 2.19, then*

$$\frac{1}{M} \cdot \sum_{(s, T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s, T, D)} \mathbf{p}_{\mathbf{g}}(s, T) = k^{k-2} + O\left(\sqrt{\frac{n}{M}}\right).$$

Proof. This is Theorem 2.16, when $a = 0$. \square

Proof of Remark 2.20. Again, Lemma 3.7 implies that for a graph G in \mathbb{G}^D ,

$$P\left(\tilde{\mathcal{G}}(D) = G\right) = \frac{\prod_{i=1}^n r_i^{d_i}}{\prod_{i,j} (1 + r_i r_j)} = e^{-H_1(\tilde{\mathbf{p}})}$$

that is independent of the choice of the graph G . Therefore, conditioning on the event $\{\tilde{\mathcal{G}}(D) \in \mathbb{G}^D\}$, the random graph $\tilde{\mathcal{G}}(D)$ is exactly the random graph $\mathcal{G}_{\mathbf{g}}(D)$.

In addition, suppose that Conjecture 2.18 holds that $e^{-\eta n \cdot \log(n)} \leq P\left(\tilde{\mathcal{G}}(D) \in \mathbb{G}^D\right)$, for some $\eta > 0$. Therefore,

$$P(\mathcal{G}_{\mathbf{g}}(D) \in A) = P\left(A | \tilde{\mathcal{G}}(D) \in \mathbb{G}^D\right) \leq e^{\eta n \cdot \log(n)} P\left(\tilde{\mathcal{G}}(D) \in A\right),$$

where A is a subset of \mathbb{G}^D . The rest of the proof is almost identical to that of Theorem 2.11.

Lastly, the factor $O(n \log(n))$ in the above equation ultimately provides us with a better bound, whereas in the proof of Theorem 2.11, we had $\mathcal{C}(D)$, which was of order $O(n^{a-\nu-\frac{1}{2}} \cdot M \cdot \log(n))$. \square

3.4. Dense graphs (Definition 2.21).

Proof of theorem 2.23. Suppose that the sequence D satisfies the dense Erdős-Rényi condition for some positive numbers c_1 , c_2 and c_3 . Let r_i be the variables defined in Eq. (3.3). Then, Lemma 4.1 in [6] implies

$$\log(|r_i|) < c_4,$$

for $1 \leq i \leq n$, and c_4 is a number that depends on c_1 , c_2 and c_3 .

Therefore, for $\mathcal{C}_2(D)$ as in Eq. (3.7),

$$\mathcal{C}(D) = \sum_{i=1}^n d_i^a |\log(r_i)| < c_4 \sum_{i=1}^n d_i^a.$$

Using Cauchy-Schwarz, and that $d_i < n$, for $1 \leq i \leq n$, we get $\mathcal{C}_2(D) < C \sqrt{\frac{n}{M}} n^{a-\frac{1}{2}} < C n^{-\frac{1}{2}+a}$. Similarly, we get a better bound for $\mathcal{C}_1(D)$ (3.5). From here, the solution is as follows in Remark 3.14 and, hence,

$$\sum_{(s,T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s,T,D)} |\mathbf{p}_a(s,T) - \tilde{\mathbf{p}}(s,T)| \leq C_k \cdot n^{\frac{-1}{4}+(a-\frac{1}{2})},$$

where $C_k > 0$ is a constant depending on k . That completes the second part of the theorem.

As for the first part of the theorem, we can either use lemma 6.2 in [6] or an exact bound from Theorem 1.4 in [1] to show that Conjecture 2.18 holds. We consider the latter here. Again, for $1 \leq i, j \leq n$, the numbers r_i s are bounded, and so are the numbers $\tilde{p}_{ij} = \frac{r_i r_j}{1+r_i r_j}$, the entries of the maximum entropy. In addition, for some $\delta(c_4)$ and as it was discussed in Remark 2.22, we have $\delta \leq \tilde{p}_{ij} \leq 1 - \delta$, which means that the maximum entropy vector $\tilde{\mathbf{p}}$ is δ -tame. Now, Theorem 1.4 in [1] states,

$$P(\tilde{G} \in \mathbb{G}^D) = e^{-H_1(\tilde{\mathbf{p}})} |\mathbb{G}^D| \approx \frac{2}{(2\pi)^{n/2} \sqrt{Q}} e^{-\frac{\mu}{2} + \nu} \geq C e^{-\gamma n \log(n)}.$$

Look at [1] for a precise definition of the variables. But let us just note that μ and ν are constants depending on D , and bounded by δ . Also, the variable Q is the determinant of a $n \times n$ matrix with entries bounded from below and above by constants depending on δ . Using Hadamard's inequality [25] to bound the determinant, we get the lower bound that we needed, which is

$$C e^{-\eta n \log(n)} \leq P(\tilde{\mathcal{G}}(D) \in \mathbb{G}^D),$$

for some constants C and $\eta > 0$. The rest is similar to the proof of Remark 2.20 and 2.11. \square

3.5. Very Sparse graphs (Definition 2.24). Effectively, the proof of Theorem 2.25 is a repetition of the arguments in the proof of Theorem 2.11, although with slight changes. We start by giving the counterparts of Lemma 3.7 and Proposition 3.6. The idea is to use $\frac{d_i}{\sqrt{M}}$ as r_i , for $1 \leq i \leq n$, so $\tilde{p}_{ij} = \frac{r_i r_j}{1+r_i r_j}$ becomes $q_{ij} = \frac{d_i d_j}{M+d_i d_j}$.

Recall that \mathcal{G}_q is a random graph with independent Bernoulli random edges with parameters q_{ij} , and

$$\mathbf{p}_q(s, T) = E[\mathbf{1}_s(T, \mathcal{G}_q(D))].$$

We state some lemmas.

Lemma 3.28. *With the same notation as above,*

(1) *let $d_q(i) := \sum_{j=1, j \neq i}^n q_{ij}$, then $d_{q,i} = d_i \left(1 - O\left(\frac{d_n^2}{M}\right)\right)$, or more precisely,*

$$d_i \left(1 - \frac{2d_n^2}{M}\right) \leq d_i \left(1 - \frac{2d_i d_n}{M}\right) \leq d_{q,i} \leq d_i.$$

(2) *For a given graph G with the degree sequence $D(G) := \{\hat{d}_1(G), \dots, \hat{d}_n(G)\}$ that may differ from $D = \{d_1, \dots, d_n\}$, we get*

$$P(\mathcal{G}_q(D) = G) = \frac{\prod_{i=1}^n \left(\frac{d_i}{\sqrt{M}}\right)^{\hat{d}_i(G)}}{\prod_{1 \leq i < j \leq n} \left(1 + \frac{d_i d_j}{M}\right)}.$$

(3) *Define*

$$P_{q,\min} = \min_{G \in \mathbb{G}_a^D} P(\mathcal{G}_q(D) = G) \text{ and } P_{q,\max} = \max_{G \in \mathbb{G}_a^D} P(\mathcal{G}_q(D) = G),$$

then,

$$\left| \log \left(\frac{P_{q,\max}}{P_{q,\min}} \right) \right| \leq 2 \log(n) \sum_{i=1}^n (d_{i,q})^a.$$

Moreover, $\sum_{i=1}^n d_i^a \leq \left(\frac{M}{n}\right)^{\frac{1}{2}} n^{(a-\frac{1}{2})}$.

Proof. (1) First, consider the following expression for $d_i - d_q(i)$,

$$\begin{aligned} d_i - \sum_{j \neq i, j \in [n]} q_{ij} &= \frac{d_i}{M} \sum_{j=1}^n d_j - \sum_{j=1, j \neq i}^n \frac{d_i d_j}{M + d_i d_j} \\ &= \frac{d_i^2}{M} + \sum_{j=1, j \neq i}^n \frac{d_i^2 d_j^2}{M(M + d_i d_j)} \\ &< \frac{d_i^2}{M} \left(1 + \sum_{j \neq i} \frac{d_n d_j}{M} \right) \\ &< d_i^2 \frac{2d_n}{M} < d_i \frac{2d_n^2}{M}. \end{aligned}$$

Second, note that the difference is positive as it is shown in the second step of the above equation.

(2) This is part 2 of Lemma 3.7, when r_i s are replaced by $\frac{d_i}{\sqrt{M}}$ s.

- (3) The proof follows from the second part of this lemma, Eq. (3.8) and (3.7), and the fact that $\log(\frac{d_i}{\sqrt{M}})$ is bounded by $\log(n)$. \square

Lemma 3.29. *There exists $\eta > 0$ such that, for large enough n ,*

$$-\eta \left(n \log(n) + \frac{d_n^4}{M} \right) \leq \log \left(P \left(\mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}^D \right) \right) \leq \log \left(P \left(\mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}_a^D \right) \right).$$

Proof. It follows from the part two of the previous lemma that,

$$(3.38) \quad P \left(\mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}^D \right) = |\mathbb{G}^D| \cdot e^L,$$

where

$$L = \sum_{i=1}^n d_i \left(\ln(d_i) - \frac{1}{2} \ln(M) \right) - \sum_{1 \leq i < j \leq n} \ln \left(1 + \frac{d_i d_j}{M} \right),$$

and also, $L = P(\mathcal{G}_{\mathbf{q}}(D) = G)$ for any graph $G \in \mathbb{G}^D$.

Next, we use Theorem 4.6 in [26], which gives the number of graphs with a given degree sequence D . Hence,

$$(3.39) \quad P \left(\mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}^D \right) = \frac{M! \exp \left(-\lambda - \lambda^2 + O\left(\frac{d_n^2}{M}\right) \right)}{\left(\frac{M}{2}\right)! 2^{\left(\frac{M}{2}\right)} \prod_{i=1}^n d_i!} \cdot e^L,$$

where λ is $\frac{1}{M} \sum_{i=1}^n \binom{d_i}{2}$. We put Eqs. (3.38) and (3.39) together, and use the stirling estimate that is

$$\log(n!) - n \log(n) + n - \frac{1}{2} \log(2\pi n) \leq \frac{c}{n},$$

where $c > 0$.

Thus, for $M = \sum_{i=1}^n d_i$ and $d_i \geq 1$,

$$(3.40) \quad \begin{aligned} & \log \left(P \left(\mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}^D \right) \right) \\ & \geq L - \lambda - \lambda^2 + O \left(\frac{d_n^2}{M} \right) + M \log(M) - M + \frac{1}{2} \log(2\pi M) \\ & \quad - \left(\frac{M}{2} \log \left(\frac{M}{2} \right) - \frac{M}{2} + \frac{1}{2} \log \left(2\pi \frac{M}{2} \right) + \frac{c}{M} \right) - \frac{M}{2} \log(2) \\ & \quad - \sum_{i=1}^n \left(d_i \log(d_i) - d_i + \frac{1}{2} \log(2\pi d_i) + \frac{c}{d_i} \right) \\ & \geq - \sum_{1 \leq i < j \leq n} \ln \left(1 + \frac{d_i d_j}{M} \right) - \lambda - \lambda^2 + O \left(\frac{d_n^2}{M} \right) + \frac{M}{2} + \frac{1}{2} \log(2) \\ & \quad - \frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^n \log(d_i) - c(n+1). \end{aligned}$$

It is time for the Taylor series for $\log(1 + \frac{d_i d_j}{M})$, which is possible because $\frac{d_i d_j}{M} < \frac{d_n^2}{M} < 1$. So, for $\lambda = \frac{1}{2M} \sum_{i=1}^n (d_i^2 - d_i) = \frac{1}{2M} \sum_{i=1}^n d_i^2 - \frac{1}{2}$,

$$\begin{aligned}
(3.41) \quad I &:= \sum_{1 \leq i < j \leq n} \ln \left(1 + \frac{d_i d_j}{M} \right) + \lambda + \lambda^2 \\
&\leq \sum_{1 \leq i < j \leq n} \left(\frac{d_i d_j}{M} - \frac{1}{2} \left(\frac{d_i d_j}{M} \right)^2 + \frac{1}{3} \left(\frac{d_i d_j}{M} \right)^3 \right) + \frac{1}{2M} \sum_{i=1}^n d_i^2 - \frac{1}{2} \\
&\quad + \left(\frac{1}{2M} \sum_{i=1}^n d_i^2 - \frac{1}{2} \right)^2 \\
&= \frac{1}{2M} \left(\sum_{i=1}^n d_i \right)^2 - \frac{1}{2} + \sum_{1 \leq i < j \leq n} \left(-\frac{1}{2M^2} d_i^2 d_j^2 + \frac{1}{3M^3} d_i^3 d_j^3 \right) \\
&\quad + \frac{1}{4M^2} \sum_{i=1}^n d_i^4 + \frac{1}{2M^2} \sum_{1 \leq i < j \leq n} d_i^2 d_j^2 - \frac{1}{2M} \sum_{i=1}^n d_i^2 + \frac{1}{4} \\
&= \frac{M}{2} - \frac{1}{4} + \frac{1}{3M^3} \sum_{1 \leq i < j \leq n} d_i^3 d_j^3 + \frac{1}{4M^2} \sum_{i=1}^n d_i^4 - \frac{1}{2M} \sum_{i=1}^n d_i^2 \\
&\leq \frac{M}{2} + \frac{d_n^4}{3M^3} \sum_{1 \leq i < j \leq n} d_i d_j + \frac{d_n^3}{4M^2} \sum_{i=1}^n d_i \\
&\leq \frac{M}{2} + \frac{d_n^4}{M}.
\end{aligned}$$

Combining (3.40) and (3.41), we get

$$\log(P(\mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}^D)) \geq -\frac{d_n^4}{M} - \frac{n}{2} \log(2\pi) - \frac{1}{2} n \log(n) - c(n+1),$$

where we used $d_i < n$. That concludes the lemma. \square

Next, we see the sparse version of Theorem 2.16, which follows from the proof of Theorem 2.25.

Lemma 3.30. *The sum of variables in Theorem 2.25 is nearly constant, i.e.*

$$\frac{1}{M} \cdot \sum_{(s,T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s,T,D)} \mathbf{p}_{\mathbf{q}}(s,T) = k^{k-2} + O\left(\sqrt{\frac{n}{M}}\right) + O\left(\frac{d_n^2}{M}\right),$$

where the constant in the O notation may depend on k .

Proof. Let $D_{\mathbf{q}}$ be the vector $(d_{\mathbf{q}}(i))$, where $1 \leq i \leq n$, and $M_{\mathbf{q}} := \sum_{i=1}^n d_{\mathbf{q}}(i)$. Replacing \tilde{p}_{ij} with q_{ij} in Theorem 3.27, we get

$$\frac{1}{M_{\mathbf{q}}} \cdot \sum_{(s,T) \in \mathbb{S}^k \times \mathbb{T}^k} \frac{1}{\psi(s,T,D_{\mathbf{q}})} \mathbf{p}_{\mathbf{q}}(s,T) = k^{k-2} + O\left(\sqrt{\frac{n}{M_{\mathbf{q}}}}\right).$$

Recall that

$$\psi(s,T,D) = \prod_{u \in V(T)} d_{s(u)}^{b_u-1},$$

where $\mathcal{V}(T)$ is the vertex set of $T \in \mathbb{T}^k$, and b_u is the degree of a vertex u in $\mathcal{V}(T)$. Thus, the first part of Lemma 3.28 demonstrates that

$$1 \leq \frac{\psi(s, T, D)}{\psi(s, T, D_{\mathbf{q}})} \leq \left(1 - \frac{2d_n^2}{M}\right)^{-k} \leq 1 + 2k \cdot \frac{d_n^2}{M},$$

and that $\left(1 - \frac{2d_n^2}{M}\right) \leq \frac{M_q}{M} \leq 1$. The combination of the above equations concludes this lemma. \square

Lemma 3.31. *We let A be any subset of $\mathbb{S}_n^k \times \mathbb{T}^k$. Then, for $\epsilon \ll 1$,*

$$P \left(\left| \frac{1}{M} \cdot \sum_{(s, T) \in A} \frac{1}{\psi(s, T, D)} (\mathbf{p}_{\mathbf{q}}(s, T) - \mathbf{1}_s(T, \mathcal{G}_{\mathbf{q}}(D))) \right| > \mu \epsilon \right) \leq e^{-(c\mu M)\epsilon^2},$$

where

$$\mu := \frac{1}{M} \cdot \sum_{(s, T) \in A} \frac{1}{\psi(s, T, D)} \mathbf{p}_{\mathbf{q}}(s, T).$$

Proof. We use the estimate $\frac{d_{\mathbf{q}}(i)}{d_i} = \Theta(1 - \frac{2d_n^2}{M})$ from part 1 of Lemma 3.28 to interchange between d_i and $d_{\mathbf{q}}(i)$. Other than that the proof is a repetition of Lemma 3.3, which we skip. \square

Proof of Theorem 2.25. We know that $d_n^2 = o(M)$, hence $d_n^2 < \frac{1}{2}M$ eventually. The first part follows from Remark 2.13 and 2.11.

Next, we show that

$$(3.42) \quad \begin{aligned} I &= \sum_{(s, T) \in \mathbb{T}^k \times \mathbb{S}^k} \frac{1}{\psi(s, T, D)} |\mathbf{p}_{\mathbf{q}}(s, T) - \mathbf{p}_{\mathbf{a}}(s, T)| \\ &\leq C_k \cdot \left(\left(\frac{n}{M} \right)^{1/4} n^{a_1} + \frac{d_n^2}{M} \right), \end{aligned}$$

where $a_1 := (a - \frac{1}{2})$ and $C_k > 0$ are constants. We combine Lemma 3.30 and Theorem 2.16 as usual. We get an equation related to Eq. (3.21) that is

$$I = 2 \left[\sum_{(s, T) \in A^-} \frac{1}{\psi(s, T, D)} (\mathbf{p}_{\mathbf{q}}(s, T) - \mathbf{p}_{\mathbf{a}}(s, T)) \right] + O \left(\left(\frac{n}{M} \right)^{1/2} n^{a_1} + \frac{d_n^2}{M} \right),$$

where

$$A^- = \left\{ (s, T) \in \mathbb{S}_n^k \times \mathbb{T}^k \mid \mathbf{p}_{\mathbf{q}}(s, T) - \mathbf{p}_{\mathbf{a}}(s, T) > 0 \right\}.$$

Now, we follow the streamline in the proof of Theorem 2.11. Much like (3.22) and without a loss of generality, we assume that the variable

$$\mu_{\mathbf{q}}^- := \sum_{(s, T) \in A^-} \frac{1}{\psi(s, T, D)} \mathbf{p}_{\mathbf{q}}(s, T)$$

is greater or equal to $\left(\frac{n}{M} \right)^{1/2} n^{3a_1}$. In addition, we let ϵ satisfy

$$\epsilon^2 = \frac{1}{\mu_{\mathbf{q}}^-} \left(\frac{n}{M} \right)^{1/2} n^{2a_1},$$

which resembles Eq. (3.22) with $n^{-\nu} = \sqrt{\frac{n}{M}}$.

We continue and combine part 3 and 4 of Lemma 3.28, and Lemma 3.31 to get,

L^-

$$\begin{aligned} &:= P \left(\left| \sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} (\mathbf{p}_{\mathbf{q}}(s,T) - \mathbf{1}_s(T, \mathcal{G}_{\mathbf{a}}(a,D))) \right| > \mu^- \epsilon \right) \\ &\leq \frac{\exp \left(O \left(\left(\frac{M}{n} \right)^{\frac{1}{2}} n^{a_1} \right) \right)}{P(\mathcal{G}_{\mathbf{q}} \in \mathbb{G}_a^D)} P \left(\left| \sum_{(s,T) \in A^-} \frac{1}{\psi(s,T,D)} (\mathbf{p}_{\mathbf{q}}(s,T) - \mathbf{1}_s(T, \mathcal{G}_{\mathbf{q}})) \right| > \mu^- \epsilon \right) \\ &\leq \exp \left(- (c\mu M) \epsilon^2 + O \left(\left(\frac{M}{n} \right)^{\frac{1}{2}} n^{a_1} \right) + O(n \log(n)) \right), \end{aligned}$$

where $\mathcal{G}_{\mathbf{q}} = \mathcal{G}_{\mathbf{q}}(D)$. The rest is the same process as in the proof of Theorem 2.11, and we end up with the bound in Eq. (3.42), i.e.

$$I \leq C_k \cdot \left(\left(\frac{n}{M} \right)^{1/4} n^{a_1} + \frac{d_n^2}{M} \right).$$

For the last parts of our theorem, we use part 1 of Lemma 3.28. Therefore,

$$\begin{aligned} (3.43) \quad P(\mathcal{G}_{\mathbf{q}}(D) \in A) &= P(\mathcal{G}_{\mathbf{q}}(D) \in A \mid \mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}^D) \\ &\leq \frac{1}{P(\mathcal{G}_{\mathbf{q}}(D) \in \mathbb{G}^D)} P(\mathcal{G}_{\mathbf{q}}(D) \in A), \end{aligned}$$

where A is a subset of \mathbb{G}^D , and again, $\mathcal{G}_{\mathbf{g}}$ is the random graph chosen uniformly from \mathbb{G}^D . Using Eq. 3.43, the last part of Lemma 3.28, and an argument identical to the proof of Remark 2.20, we obtain

$$(3.44) \quad \sum_{(s,T) \in \mathbb{S}_n^k \times \mathbb{T}^k} \frac{1}{\psi(s,T,D)} |\mathbf{p}_{\mathbf{q}}(s,T) - \mathbf{p}_{\mathbf{g}}(s,T)| \leq C_k \cdot \left(\left(\frac{n \log(n)}{M} \right)^{1/2} + \frac{d_n^2}{M} \right),$$

where

$$\mathbf{p}_{\mathbf{g}}(s,T) = E[\mathbf{1}_s(T, \mathcal{G}_{\mathbf{g}})],$$

as in Conjecture 2.19.

Finally, parts 2 and 3 of the theorem follow from the first part of the theorem, Eq. (3.42) and (3.44), and the triangle inequality. \square

Remark 3.32. Although the bound for the differences $|\mathbf{p}_{\mathbf{g}}(s,T) - \tilde{\mathbf{p}}(s,T)|$ in the second part of the theorem is $o(1)$, compared to Conjecture 2.19, it is not optimal. The reason is that we do not get the lower bound in Conjecture 2.18. Instead, we use parameters $\mathbf{p}_{\mathbf{q}}(s,T)$ and $\mathbf{p}_{\mathbf{a}}(s,T)$ as a middle step to achieve our bound.

3.6. Bipartite graphs (proof of Theorem 2.30). Although the setup is a little bit different here, the proof operates along similar lines as the proof of Remark 2.20. The main difference is that everything splits into two sets of variables. For example, there is a related version of Eq. (3.4) for the maximum entropy $\tilde{\mathbf{p}} \in \mathbb{P}^{D_1, D_2}$. We write, for $1 \leq i \leq n_1$ and $1 \leq j \leq n_2$,

$$\tilde{p}_{ij} = \frac{r_{1,i} r_{2,j}}{1 + r_{1,i} r_{2,j}},$$

where $x^* = (r_{1,1}, \dots, r_{1,n_1})$ and $y^* = (r_{2,1}, \dots, r_{2,n_2})$ are two positive vectors.

In regard to the ordered trees, we restrict our sums to the trees $(s, T) \in \mathbb{S}_n^k \times \mathbb{T}^k$ that $s(T)$ does not have any edge with both ends in vertices of either part 1 or part 2. We let \mathcal{T}_b^k be the set of such trees. We check that

$$(3.45) \quad \frac{1}{M} \cdot \sum_{(s,T) \in \mathcal{T}_b^k} \frac{1}{\psi(s, T, D)} \mathbf{P}_b(s, T) = k^{k-2} + O\left(\sqrt{\frac{n}{M}}\right)$$

is still valid. Although it sounds contradictory to Theorem 3.27 since \mathcal{T}_b^k is a subset of $\mathbb{S}_n^k \times \mathbb{T}^k$, we note that the definitions of $d_{1,i}$ and $d_{2,i}$ are different from d_i in Theorem 3.27. Here,

$$d_{1,i} := \sum_{j=1}^{n_1} \tilde{p}_{ij}, \text{ and } d_{2,i} := \sum_{j=1}^{n_2} \tilde{p}_{ij},$$

as opposed to

$$d_i := \sum_{j=1}^n \tilde{p}_{ij},$$

where $n = n_1 + n_2$.

Next, Theorem 1-1 of [1] gives the following bounds

$$(3.46) \quad (n_1 n_2)^{-\eta(n_1+n_2)} \leq P(\mathcal{G}_b(D_1, D_2) \in \mathbb{G}^{D_1, D_2}) = e^{-H_2(\tilde{\mathbf{P}})} |\mathbb{G}^{D_1, D_2}|,$$

for some positive η independent of n . The above term $(n_1 n_2)^{-\eta(n_1+n_2)}$ is bounded by $e^{-\eta n \log(n^2)}$. Equations (3.45) and (3.46) are enough to produce a proof using the same method in the proof of Remark 2.20, and we skip the details.

Remark 3.33. The Theorem 1-1 of [1] only requires that the polytope \mathbb{P}^{D_1, D_2} has a non-empty interior. That gives a proof of Theorem 2.30 without any extra conditions on the degree sequence like part 2 and 3 of Assumption 2.7. We also believe that one can prove all the previous results without any extra condition on the degree sequence.

4. ACKNOWLEDGEMENTS

The author would like to thank S. R. Srinivasa Varadhan for reviewing the manuscript and his useful comments. Thanks to S. Chatterjee for suggesting the problem. Thanks to A. Barvinok, M. Harel, A. Krishnan, and A. Munez for fruitful discussions and references.

APPENDIX A. A CONCENTRATION INEQUALITY.

We also need the following concentration theorem for the proof of Theorem 2.11 that is inspired by a paper by Janson [21]. This is the generalized version of Theorem 1 from Janson's paper, and we modified the proof for our purpose. Therefore, we begin this part with some notations and a theorem.

Consider a set $\{J_i\}_{i \in Q}$ of independent random indicator variables and a family $\{Q(\alpha)\}_{\alpha \in A}$ of subsets of the index set Q , and define $\mathbf{1}_\alpha = \prod_{i \in Q(\alpha)} J_i$ and $S = \sum_{\alpha \in A} \frac{1}{\omega_\alpha} \mathbf{1}_\alpha$, where ω_α are positive numbers. [In other words, S counts the weighted number of the given sets $Q\{\alpha\}$ that are contained in the random set $\{i \in Q : J_i = 1\}$, with independently appearing elements.] We assume, for the sake

of simplicity, that the index set A is finite, but it is easy to see that the results extend to infinite sums, provided $E[S] < \infty$.

Write $\alpha \sim \beta$ if $Q(\alpha) \cap Q(\beta) \neq \emptyset$ but $\alpha \neq \beta$, and define

$$\begin{aligned} p_\alpha &= E[\mathbf{1}_\alpha], \\ \lambda &= E[S] = \sum_\alpha \frac{1}{\omega_\alpha} p_\alpha, \\ \delta_1 &= \frac{1}{\lambda} \sum_\alpha \frac{p_\alpha}{\omega_\alpha^2}, \\ \delta_2 &= \frac{1}{\lambda} \sum_\alpha \sum_{\beta \sim \alpha} \frac{1}{\omega_\alpha \omega_\beta} E[\mathbf{1}_\alpha \mathbf{1}_\beta]. \end{aligned}$$

Theorem A.1. *With notation above and $0 \leq \epsilon \leq 1$ then*

$$P(S \leq (1 - \epsilon)\lambda) \leq \exp \left[-\frac{\lambda}{\delta_1 + \delta_2} (\epsilon + (1 - \epsilon) \log(1 - \epsilon)) \right].$$

We want to use the Chernoff bound, but first we need an upper bound for the moment-generating function. So the proof of Theorem A.1 follows our next lemma.

Lemma A.2. *Using the preceding notations in Theorem A.1 and for $t \geq 0$, we have*

$$E[e^{-tS}] \leq \exp \left[-\frac{\lambda}{\delta_1 + \delta_2} (1 - e^{-(\delta_1 + \delta_2)t}) \right].$$

Proof. Let $\psi(t) = E[e^{-tS}]$, for $t \geq 0$. Then

$$-\frac{d\psi(t)}{dt} = E[Se^{-tS}] = \sum_\alpha E \left[\frac{1}{\omega_\alpha} \mathbf{1}_\alpha e^{-tS} \right].$$

We split S into two parts; the part that is dependent on $\mathbf{1}_\alpha$: $S'_\alpha = \frac{1}{\omega_\alpha} \mathbf{1}_\alpha + \sum_{\alpha \sim \beta} \frac{1}{\omega_\beta} \mathbf{1}_\beta$, and $S''_\alpha = S - S'_\alpha$, which is independent of $\mathbf{1}_\alpha$. Thus,

$$E[\mathbf{1}_\alpha e^{-tS}] = p_\alpha E[e^{-tS'_\alpha - tS''_\alpha} | \mathbf{1}_\alpha = 1].$$

The event $\mathbf{1}_\alpha = 1$ fixes $J_i : i \in Q(\alpha)$. Since $e^{-tS'_\alpha}$ and $e^{-tS''_\alpha}$ are decreasing functions of the remaining $J_i : i \in Q$, using the FKG inequality we get

$$\begin{aligned} E[\mathbf{1}_\alpha e^{-tS}] &\geq p_\alpha E[e^{-tS'_\alpha} | \mathbf{1}_\alpha = 1] E[e^{-tS''_\alpha} | \mathbf{1}_\alpha = 1] \\ &= p_\alpha E[e^{-tS'_\alpha} | \mathbf{1}_\alpha = 1] E[e^{-tS''_\alpha}] \\ (A.1) \quad &\geq p_\alpha E[e^{-tS'_\alpha} | \mathbf{1}_\alpha = 1] \psi(t). \end{aligned}$$

Now summing over α and using Jensen's inequality twice we have

$$\begin{aligned}
-\frac{d \log(\psi(t))}{dt} &= \frac{1}{\psi(t)} \sum_{\alpha} \frac{1}{\omega_{\alpha}} E[\mathbf{1}_{\alpha} e^{-tS}] \geq \sum_{\alpha} \frac{p_{\alpha}}{\omega_{\alpha}} E[e^{-tS'_{\alpha}} | \mathbf{1}_{\alpha} = 1] \\
&\geq \sum_{\alpha} \frac{p_{\alpha}}{\omega_{\alpha}} \exp[-tE[S'_{\alpha} | \mathbf{1}_{\alpha} = 1]] \\
&= \lambda \sum_{\alpha} \frac{1}{\lambda} \frac{p_{\alpha}}{\omega_{\alpha}} \exp[-tE[S'_{\alpha} | \mathbf{1}_{\alpha} = 1]] \\
&\geq \lambda \exp \left[-t \left(\sum_{\alpha} \frac{1}{\lambda} \frac{p_{\alpha}}{\omega_{\alpha}} E[S'_{\alpha} | \mathbf{1}_{\alpha} = 1] \right) \right] \\
&= \lambda \exp \left[-\frac{t}{\lambda} \left(\sum_{\alpha} \frac{1}{\omega_{\alpha}} E[S'_{\alpha} \mathbf{1}_{\alpha}] \right) \right] \\
&= \lambda \exp \left[-\frac{t}{\lambda} \left(\sum_{\alpha} \frac{1}{\omega_{\alpha}^2} E[\mathbf{1}_{\alpha}^2] + \sum_{\alpha \sim \beta} \frac{1}{\omega_{\alpha} \omega_{\beta}} E[\mathbf{1}_{\alpha} \mathbf{1}_{\beta}] \right) \right] \\
\text{(A.2)} \quad &= \lambda \exp[-t(\delta_1 + \delta_2)].
\end{aligned}$$

Therefore, $(\psi(0) = 1)$

$$-\log(\psi(t)) \geq \int_0^t \lambda e^{-t(\delta_1 + \delta_2)} = \frac{\lambda}{\delta_1 + \delta_2} (1 - e^{-(\delta_1 + \delta_2)t}).$$

□

Proof of Theorem A.1. Now we are ready to use Chernoff's bound,

$$P(S \leq (1 - \epsilon)\lambda) \leq e^{t(1 - \epsilon)\lambda} E[e^{-tS}] \leq \exp \left[t(1 - \epsilon)\lambda - \frac{\lambda}{\delta_1 + \delta_2} (1 - e^{-(\delta_1 + \delta_2)t}) \right].$$

Optimizing over t , we get $t = -(\delta_1 + \delta_2)^{-1} \log(1 - \epsilon)$. Thus,

$$\begin{aligned}
P(S \leq (1 - \epsilon)\lambda) &\leq \exp \left[-(1 - \epsilon) \log(1 - \epsilon) \frac{\lambda}{\delta_1 + \delta_2} - \frac{\lambda}{\delta_1 + \delta_2} \epsilon \right] \\
&= \exp \left[-\frac{\lambda}{\delta_1 + \delta_2} [\epsilon + (1 - \epsilon) \log(1 - \epsilon)] \right].
\end{aligned}$$

This completes the proof. □

APPENDIX B. REGULARITY OF r_i S.

Recall that the vector $(\tilde{p}_{ij})_{1 \leq i \neq j \leq n} \in \mathbb{P}^D$ is the minimizer of

$$H_1(x) = \sum_{i < j} H(x_{ij}), \text{ where } H(x) = -x \ln(x) - (1 - x) \ln(1 - x).$$

In addition, we have

$$\text{(B.1)} \quad d_i = \sum_{j \in [n] \setminus \{i\}} \tilde{p}_{ij},$$

and as in Eq. 2.1, $\tilde{p}_{ij} = \frac{r_i r_j}{1 + r_i r_j}$ that r_i s are positive numbers, and $1 \leq i \leq n$.

Lemma B.1. *Suppose that $d_1 \leq \dots \leq d_n$, then,*

- a) $r_1 \leq \dots \leq r_n$,
- b) and $r_1 r_n > \frac{1}{n}$.
- c) If $r_k \geq 1$, for some $1 \leq k \leq n$, then $r_{k+1}/r_k < n^4$.
- d) If $r_k > n^2$, for some $1 \leq k \leq n$, then $\sum d_i \leq \frac{1}{2}M$, where the sum is over $1 \leq i \leq n - d_k - 1$.

Proof. a) We observe that

$$\begin{aligned} d_j - d_i &= \sum_{k \neq i, j} \frac{r_j r_k}{1 + r_j r_k} - \frac{r_i r_k}{1 + r_i r_k} \\ &= (r_j - r_i) \sum_{k \neq i, j} \frac{r_k}{(1 + r_i r_k)(1 + r_j r_k)}. \end{aligned}$$

The r_i s are positive, as well as the last sum in the above equation. Hence, the terms $d_j - d_i$ and $r_j - r_i$ have the same sign, and this finishes part a.

- b) Let us see that $\tilde{p}_{ij} = \frac{r_i r_j}{1 + r_i r_j}$ are increasing both in i and j , because r_i s are positive and are increasing by part a, and also $f(x) = \frac{x}{1+x}$ is increasing in x , for $x \geq 0$. therefore, (B.1) implies

$$1 \leq d_1 = \sum_{j=2}^n \frac{r_1 r_j}{1 + r_1 r_j} \leq (n-1) \frac{r_1 r_n}{1 + r_1 r_n} \leq n(r_1 r_n).$$

That is what we want.

- c) We prove the problem using contradiction. So, suppose $r_k \geq 1$ and $r_{k+1}/r_k \geq n^4$. We define $I = \{i | r_{k+1} \leq r_i\}$ and $J = \{j | r_j \leq \frac{n^2}{r_{k+1}}\}$. Therefore,

$$0 < \frac{r_j r_l}{1 + r_j r_l} \leq \frac{\frac{n^2}{r_{k+1}} r_k}{1 + \frac{n^2}{r_{k+1}} r_k} < \frac{1}{n^2},$$

for $j \in J$, and $l \in [n] \setminus I$. In addition,

$$1 > \frac{r_i r_l}{1 + r_i r_l} > \frac{r_{k+1} \frac{n^2}{r_{k+1}}}{1 + r_{k+1} \frac{n^2}{r_{k+1}}} > 1 - \frac{1}{n^2},$$

for $i \in I$ and $l \in [n] \setminus J$. We observe that, for the number $U := \sum_{i \in I} d_i - \sum_{j \in J} d_j$,

$$\begin{aligned} U &= \sum_{i \in I} \sum_{l \in [n] \setminus \{i\}} \frac{r_i r_l}{1 + r_i r_l} - \sum_{j \in J} \sum_{l \in [n] \setminus \{j\}} \frac{r_j r_l}{1 + r_j r_l} \\ &= \sum_{i \in I} \sum_{l \in [n] \setminus J \cup \{i\}} \frac{r_i r_l}{1 + r_i r_l} - \sum_{j \in J} \sum_{l \in [n] \setminus I \cup \{j\}} \frac{r_j r_l}{1 + r_j r_l} \\ &\geq |I| \cdot (n - |J| - 1) \left(1 - \frac{1}{n^2}\right) \\ &> |I| \cdot (n - |J| - 1) - 1. \end{aligned}$$

Moreover,

$$|I| \cdot (n - |J| - 1) - 1 < U \leq \sum_{i \in I} \sum_{l \in [n] \setminus J \cup \{i\}} \frac{r_i r_l}{1 + r_i r_l} < |I| \cdot (n - |J| - 1),$$

which is impossible since U is an integer.

d) By part a, r_i s are increasing in i . Let $I = \{i \mid r_i < n^{-1}\}$, then

$$\begin{aligned} d_k &= \sum_{i \in [n] \setminus \{k\}} \frac{r_i r_k}{1 + r_i r_k} \geq \sum_{i \notin I \cup \{k\}} \frac{r_i r_k}{1 + r_i r_k} \\ &\geq (n - |I| - 1) \frac{n}{1 + n} \\ &> n - |I| - 2, \end{aligned}$$

since $r_k \geq n^2$. We note that d_k is an integer, so $d_k \geq n - |I| - 1$. Now, there are at most n^2 pairs of i and j in I , and $\frac{r_i r_j}{1 + r_i r_j} \leq \frac{1}{n^2 + 1}$. That implies

$$\begin{aligned} \sum_{i \in I} d_i &= \sum_{i \in I} \left(\sum_{j \in I, j \neq i} + \sum_{j \notin I} \right) \frac{r_i r_j}{1 + r_i r_j} \\ &< \frac{n^2}{n^2 + 1} + \sum_{i \in I} \sum_{j \notin I} \frac{r_j r_i}{1 + r_j r_i} \\ &\leq \frac{n^2}{n^2 + 1} + \sum_{j \notin I} d_j \\ &< 1 + M - \sum_{i \in I} d_i. \end{aligned}$$

In addition, $\sum_{i \in I} d_i$ is an integer, so $\sum_{i \in I} d_i \leq \frac{M}{2}$. Ultimately, we close this lemma by $\sum_{i=1}^{n-d_k-1} d_i \leq \sum_{i=1}^{|I|} d_i \leq \frac{M}{2}$. \square

Recall that $\tilde{\mathcal{G}}(D)$ is a random graph with independent Bernoulli random edges with parameters \tilde{p}_{ij} .

Lemma B.2. *The following variational problems are equivalent,*

$$\inf_{\vec{x} \in \mathbb{R}^n} F(\vec{x}) = \inf_{\vec{x} \in (\mathbb{R}^{>0})^n} G(\vec{x}) = \sup_{p \in \mathbb{P}^D} H_1(p),$$

where

$$F(\vec{x}) = - \sum_{i=1}^n d_i x_i + \sum_{i < j} \log(1 + e^{x_i + x_j}),$$

and

$$G(\vec{x}) = - \sum_{i=1}^n d_i \log(x_i) + \sum_{i < j} \log(1 + x_i x_j).$$

In addition, the supremum of H_1 is equal to $-\log \left(P \left(\tilde{G} = G \right) \right)$ for any graph G with a degree sequence that is equal to D .

Proof. In the proof of Proposition 2.1, we saw that $H_1(x)$ takes its maximum $(\tilde{p}_{ij})_{1 \leq i \neq j \leq n}$ in the interior of \mathbb{P}^D . In regard to the function $F(\vec{x})$, it is strictly convex and, hence, has at most one minimum. Actually, the minimum is $\vec{\lambda} := (\log(r_1), \dots, \log(r_n))$, since the gradient of $F(\vec{x})$ at $\vec{\lambda}$ is

$$\partial_i F(\vec{\lambda}) = -d_i + \sum_{j \neq i} \frac{e^{\lambda_i + \lambda_j}}{1 + e^{\lambda_i + \lambda_j}} = -d_i + \sum_{j \neq i} \frac{r_i r_j}{1 + r_i r_j} = 0.$$

Thus, $\vec{\lambda}$ is a critical point and the unique minimum of $F(\vec{x})$. In addition, by a change of variable we get $G(\vec{x})$ from $F(\vec{x})$. So, $\vec{r} = (r_1, \dots, r_n)$ solves the infimum problem for function $G(\vec{x})$, or $\inf_{\vec{x} \in (\mathbb{R}_{>0})^n} G(\vec{x}) = G(\vec{r})$.

Next, we rewrite $F(\vec{\lambda}) = G(\vec{r})$ in terms of \tilde{p}_{ij} ,

$$\begin{aligned}
G(\vec{r}) &= - \sum_{i=1}^n d_i \log(r_i) + \sum_{i<j} \log(1 + r_i r_j) \\
&= - \sum_{i=1}^n \sum_{j \neq i} \frac{r_i r_j}{1 + r_i r_j} \log(r_i) + \sum_{i<j} \log(1 + r_i r_j) \\
&= - \sum_{i<j} \frac{r_i r_j}{1 + r_i r_j} [\log(r_i) + \log(r_j) - \log(1 + r_i r_j)] \\
&\quad + \frac{1}{1 + r_i r_j} \log(1 + r_i r_j) \\
&= - \sum_{i<j} \tilde{p}_{ij} \log(\tilde{p}_{ij}) - (1 - \tilde{p}_{ij}) \log(1 - \tilde{p}_{ij}) \\
&= H_1(\tilde{\mathbf{p}}).
\end{aligned}$$

This completes the first part of the lemma.

On account of the vector D satisfying the strict Erdős- Gallai conditions (2.2), there exists a graph G with the degree sequence $D(G)$ that is equal to D . Let us use Lemma 3.5 with the graph G , and the above equation to reach

$$\begin{aligned}
-\log \left(P \left(\tilde{\mathcal{G}}(D) = G \right) \right) &= -\log \frac{\prod_{i=1}^n r_i^{d_i}}{\prod_{i,j} (1 + r_i r_j)} \\
&= - \sum_{i=1}^n d_i \log(r_i) + \sum_{i<j} \log(1 + r_i r_j) \\
&= G(\vec{r}).
\end{aligned}$$

□

Lemma B.3. *If $M = \sum_{i=1}^n d_i \leq \binom{n}{2}$, then for $G \in \mathbb{G}^D$,*

$$M \cdot \log \left(\frac{M}{n(n-1)} \right) \leq \log \left(P \left(\tilde{\mathcal{G}}(D) = G \right) \right).$$

Proof. First, the previous lemma provides that $-\log \left(P \left(\tilde{\mathcal{G}}(D) = G \right) \right) = H_1(\tilde{\mathbf{p}})$, and moreover,

$$\sum_{1 \leq i \neq j \leq n} \tilde{p}_{ij} = \frac{1}{2} \sum_{1 \leq i < j \leq n} \tilde{p}_{ij} = \frac{M}{2}.$$

Second, the function $H(x) = -x \log(x) - (1-x) \log(1-x)$ is a concave function. So,

$$\begin{aligned}
& \frac{2}{n(n-1)} H_1(\tilde{p}) \\
&= \frac{2}{n(n-1)} \sum_{i < j} -\tilde{p}_{ij} \log(\tilde{p}_{ij}) - (1 - \tilde{p}_{ij}) \log(1 - \tilde{p}_{ij}) \\
&\leq -\frac{M}{n(n-1)} \log\left(\frac{M}{n(n-1)}\right) - \left(1 - \frac{M}{n(n-1)}\right) \log\left(1 - \frac{M}{n(n-1)}\right) \\
&\leq -\frac{2M}{n(n-1)} \log\left(\frac{M}{n(n-1)}\right),
\end{aligned}$$

and we used the inequality $x \log(x) \leq (1-x) \log(1-x)$ for $x \leq \frac{1}{2}$. \square

REFERENCES

- [1] Barvinok, A., 2010a. On the number of matrices and a random matrix with prescribed row and column sums and 0–1 entries. *Adv. Math.* 224, 316–339. URL: <http://dx.doi.org/10.1016/j.aim.2009.12.001>, doi:10.1016/j.aim.2009.12.001.
- [2] Barvinok, A., 2010b. What does a random contingency table look like? *Combin. Probab. Comput.* 19, 517–539. URL: <http://dx.doi.org/10.1017/S0963548310000039>, doi:10.1017/S0963548310000039.
- [3] Barvinok, A., 2012. Matrices with prescribed row and column sums. *Linear Algebra Appl.* 436, 820–844. URL: <http://dx.doi.org/10.1016/j.laa.2010.11.019>, doi:10.1016/j.laa.2010.11.019.
- [4] Barvinok, A., Hartigan, J.A., 2010. Maximum entropy Gaussian approximations for the number of integer points and volumes of polytopes. *Adv. in Appl. Math.* 45, 252–289. URL: <http://dx.doi.org/10.1016/j.aam.2010.01.004>, doi:10.1016/j.aam.2010.01.004.
- [5] Barvinok, A., Hartigan, J.A., 2012. An asymptotic formula for the number of non-negative integer matrices with prescribed row and column sums. *Trans. Amer. Math. Soc.* 364, 4323–4368. URL: <http://dx.doi.org/10.1090/S0002-9947-2012-05585-1>, doi:10.1090/S0002-9947-2012-05585-1.
- [6] Barvinok, A., Hartigan, J.A., 2013. The number of graphs and a random graph with a given degree sequence. *Random Structures Algorithms* 42, 301–348. URL: <http://dx.doi.org/10.1002/rsa.20409>, doi:10.1002/rsa.20409.
- [7] Bollobás, B., Riordan, O., 2009. Metrics for sparse graphs, in: *Surveys in combinatorics 2009*. Cambridge Univ. Press, Cambridge. volume 365 of *London Math. Soc. Lecture Note Ser.*, pp. 211–287.
- [8] Borgs, C., Chayes, J., Lovász, L., Sós, V.T., Vesztegombi, K., 2006. Counting graph homomorphisms, in: *Topics in discrete mathematics*. Springer, Berlin. volume 26 of *Algorithms Combin.*, pp. 315–371. URL: http://dx.doi.org/10.1007/3-540-33700-8_18, doi:10.1007/3-540-33700-8_18.

- [9] Borgs, C., Chayes, J., Lovász, L., Sós, V.T., Vesztergombi, K., 2012. Convergent sequences of dense graphs ii: Multi-way cuts and statistical physics. *Ann. Math.* , 151–219 URL: <http://annals.math.princeton.edu/wp-content/uploads/annals-v176-n1-p02-p.pdf>.
- [10] Borgs, C., Chayes, J.T., Lovász, L., Sós, V.T., Vesztergombi, K., 2008. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* 219, 1801–1851. URL: <http://dx.doi.org/10.1016/j.aim.2008.07.008>, doi:10.1016/j.aim.2008.07.008.
- [11] Canfield, E.R., Greenhill, C., McKay, B.D., 2008. Asymptotic enumeration of dense 0-1 matrices with specified line sums. *J. Combin. Theory Ser. A* 115, 32–66. URL: <http://dx.doi.org/10.1016/j.jcta.2007.03.009>, doi:10.1016/j.jcta.2007.03.009.
- [12] Canfield, E.R., McKay, B.D., 2010. Asymptotic enumeration of integer matrices with large equal row and column sums. *Combinatorica* 30, 655–680. URL: <http://dx.doi.org/10.1007/s00493-010-2426-1>, doi:10.1007/s00493-010-2426-1.
- [13] Cayley, A., 1889. A theorem on trees. *Quart. J. Math.* 23, 376–378.
- [14] Chatterjee, S., 2012. The missing log in large deviations for triangle counts. *Random Structures Algorithms* 40, 437–451. URL: <http://dx.doi.org/10.1002/rsa.20381>, doi:10.1002/rsa.20381.
- [15] Chawla, S., 2004. Randomized algorithms. URL: <http://www.cs.cmu.edu/afs/cs/academic/class/15859-f04/www/scribes/lec9.pdf>.
- [16] Erdős, P., Gallai, T., 1960. Graphen mit punkten vorgeschriebenen grades. *Mat. Lapok* 11, 264–274.
- [17] Gao, P., 2011. Distribution of certain sparse spanning subgraphs in random graphs. *Random Struct. Algorithms* URL: <http://arxiv.org/pdf/1105.5913v1>.
- [18] Gao, P., Su, Y., Wormald, N., 2012. Induced subgraphs in sparse random graphs with given degree sequences. *European J. Combin.* 33, 1142–1166. URL: <http://dx.doi.org/10.1016/j.ejc.2012.01.009>, doi:10.1016/j.ejc.2012.01.009.
- [19] Greenhill, C., McKay, B.D., 2012. Counting loopy graphs with given degrees. *Linear Algebra Appl.* 436, 901–926. URL: <http://dx.doi.org/10.1016/j.laa.2011.03.052>, doi:10.1016/j.laa.2011.03.052.
- [20] Greenhill, C., McKay, B.D., Wang, X., 2006. Asymptotic enumeration of sparse 0-1 matrices with irregular row and column sums. *J. Combin. Theory Ser. A* 113, 291–324. URL: <http://dx.doi.org/10.1016/j.jcta.2005.03.005>, doi:10.1016/j.jcta.2005.03.005.
- [21] Janson, S., 1990. Poisson approximation for large deviations. *Random Structures Algorithms* 1, 221–229. URL: <http://dx.doi.org/10.1002/rsa.3240010209>, doi:10.1002/rsa.3240010209.
- [22] Janson, S., Ruciński, A., 2011. Upper tails for counting objects in randomly induced subhypergraphs and rooted random graphs. *Ark. Mat.* 49, 79–96. URL: <http://dx.doi.org/10.1007/s11512-009-0117-1>, doi:10.1007/s11512-009-0117-1.

- [23] Lovász, L., Szegedy, B., 2006. Limits of dense graph sequences. J. Combin. Theory Ser. B 96, 933–957. URL: <http://arxiv.org/abs/math/0408173>.
- [24] Mahadev, N.V.R., Peled, U.N., 1995. Threshold graphs and related topics. volume 56 of *Annals of Discrete Mathematics*. North-Holland Publishing Co., Amsterdam.
- [25] Maz'ya, V., Shaposhnikova, T., 1998. Jacques Hadamard, a universal mathematician. volume 14 of *History of Mathematics*. American Mathematical Society, Providence, RI.
- [26] McKay, B.D., 1985. Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combin.* 19, 15–25.
- [27] McKay, B.D., 2010. Subgraphs of random graphs with specified degrees, in: *Proceedings of the International Congress of Mathematicians. Volume IV*, Hindustan Book Agency, New Delhi. pp. 2489–2501.
- [28] McKay, B.D., 2011. Subgraphs of dense random graphs with specified degrees. *Combin. Probab. Comput.* 20, 413–433. URL: <http://dx.doi.org/10.1017/S0963548311000034>, doi:10.1017/S0963548311000034.
- [29] McKay, B.D., Wormald, N.C., 1990. Asymptotic enumeration by degree sequence of graphs of high degree. *European J. Combin.* 11, 565–580.
- [30] McKay, B.D., Wormald, N.C., 1991. Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica* 11, 369–382. URL: <http://dx.doi.org/10.1007/BF01275671>, doi:10.1007/BF01275671.
- [31] Prufer, H., 1918. Neuer beweis eines satzes über permutationen. *Arch. Math. Phys.* 27, 742–744.

Courant Institute of Mathematical Sciences
 New York University
 251 Mercer Street
 New York, NY-10012
 United States of America
 mehrdad@cims.nyu.edu